

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:14:49 ; Search time 76 Seconds
(without alignments)
38.573 Million cell updates/sec

Title: US-09-929-788-1
Perfect score: 118
Sequence: 1 FIIHIFRGIYHAGRSIGRFLTG 22

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
A.GeneSeq_101002:*
1: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SID2/gcgcdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	100.0	22	23	AAE20172
2	118	100.0	22	23	ABR07720
3	87	73.7	22	23	AAE20173
4	87	73.7	22	23	ABR07721
5	84	71.2	22	23	AAE20174
6	84	71.2	22	23	ABR07722
7	59	50.0	24	23	AAE20177
8	49	41.5	22	22	ABG25467
9	49	41.5	22	22	ABG24377
10	48	40.7	509	22	AAU38243

11	48	40.7	510	19	AAW69358
12	48	40.7	510	22	AAU34573
13	48	40.7	883	23	ABW54148
14	47	39.8	512	21	AAV74582
15	47	39.8	513	21	AAV74583
16	47	39.8	1721	19	AAW52847
17	46	39.0	480	22	ABW66251
18	46	39.0	480	22	ABW70230
19	46	39.0	496	21	AAW27883
20	46	39.0	522	21	AAW27882
21	46	39.0	687	21	AAW27881
22	45	38.1	393	21	AAV77947
23	45	38.1	696	22	AAU04874
24	45	38.1	3413	19	AAW52849
25	45	38.1	5069	19	AAW52846
26	44.5	37.7	2724	22	AAU08681
27	44.5	37.7	2733	22	AAU08680
28	44	37.3	271	22	AAW82031
29	44	37.3	282	23	ABP38310
30	44	37.3	399	19	AAW84138
31	44	37.3	399	20	AAW95505
32	44	37.3	399	21	AAV92600
33	44	37.3	399	21	AAV56046
34	44	37.3	403	22	AAW31685
35	44	37.3	491	23	ABW08382
36	44	37.3	765	22	ABW28315
37	43.5	36.9	852	22	ABW59714
38	43.5	36.9	43	22	ABW64019
39	43	36.4	43	22	ABW20589
40	43	36.4	43	22	AAW68359
41	43	36.4	43	22	AAW28667
42	43	36.4	43	22	AAW03906
43	43	36.4	43	23	ABW37920
44	43	36.4	73	20	AAV02932
45	43	36.4	107	22	AAO01861

ALIGNMENTS

RESULT 1
ID AAE20172 standard; peptide: 22 AA.
AAE20172;
18-JUN-2002 (first entry)
Fish mast cell antimicrobial peptide, piscidin 3 (P3).
Antimicrobial peptide; endobiotic peptide; mast cell; piscidin; fish;
methicillin; vancomycin; streptogramin; microbial infection; stress;
bacterial disease; fish health; screening; gene therapy; antimicrobial;
transluciferase; fungicide; protozoicide; antibacterial; antiparasitic.
XX
OS Morone chrysops.
OS Morone saxatilis.
XX
PN WO200214345-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US41696.
XX
PR 15-AUG-2000; 2000US-225354P.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Noga EJ, Silphaduang U;
XX
DR WPI; 2002-269176/31.
XX
PT Novel antimicrobial peptides (endobiotic peptides) isolated from mast

PT cells, specifically fish mast cells, useful for treating microbial
PT infections in humans or animals and for reducing antibiotic resistance
XX in bacteria -
XX
PS Claim 4; Page 17; 31pp; English.
XX
CC The invention relates to antimicrobial peptide (endobiotic peptide)
CC isolated from a mast cell. Such peptides obtained from fish mast cells
CC are referred as piscidins. The antimicrobial peptide is useful for
CC treating microbial infection in a subject, and for reducing antibiotic
CC (e.g., methicillin, vancomycin or streptogramin) resistance in bacteria
CC such as *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus aureus*,
CC *Klebsiella pneumoniae*, *Streptococcus faecalis* or *Shigella flexneri*. These
CC peptides are useful for treating microbial infections in human or animals
CC and for preparing a medicament for treating microbial infections. These
CC peptides are also useful for treating infections caused by microorganisms
CC other than bacteria, e.g., fungi, yeast, protozoa, parasite, etc. These
CC peptides are useful for treating stress or bacterial disease in fish and
CC monitoring fish health. The inverse relationship between endobiotic
CC levels and stress also allows assessment of freshness of fish food
CC product. Nucleic acid sequences encoding these peptides are useful for
CC diagnostic purposes, treating stress in a fish and for screening
CC additional endobiotic peptides. They are also useful in gene therapy.
CC The present sequence is piscidin 3 (P3) peptide isolated from the gills
CC of hybrid striped bass (Morone saxatilis x Morone chrysops).
CC
SQ Sequence 22 AA:

Query Match 100.0%; Score 118; DB 23; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.8e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
Db 1 FHHIFRGIVHAGRSIGRFLTG 22
|||||
RESULT 2
AB07720 ID ABB07720 standard; peptide; 22 AA.
XX
AC ABB07720;
XX
DT 10-JUN-2002 (first entry)
XX
DE Antimicrobial peptide #1 from hybrid striped bass.
XX
KW Fish; hybrid striped bass; endobiotic; antibiotic; tranquilizer;
XX antimicrobial; antibacterial; fish food product.
XX
OS Morone saxatilis x Morone chrysops.
XX
PN WO200214346-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001MO-US41697.
XX
PR 15-AUG-2000; 2000US-225354P.
XX
XX (UYNC-) UNTV NORTH CAROLINA STATE.
XX
PI Noga EJ, Silphaduang U;
XX
XX WPI; 2002-269177/31.
XX
DR Novel antimicrobial compound (endobiotic peptide isolated from fish)
XX PT useful for monitoring fish health and monitoring freshness of fish food
XX PT product and for screening compounds which treat stress in fish -
XX
PS Claim 1; Page 16; 25pp; English.
XX
CC The invention provides antimicrobial peptides (AB07720-23) isolated from

CC hybrid striped bass. The peptides were purified from the endobiotic
CC family. The endogenous endobiotic peptides exhibiting antimicrobial
CC activity) are useful for monitoring fish health, monitoring the freshness
CC of the fish product such as fresh, refrigerated, or frozen, fish food
CC product, and for screening compounds useful for treating stress in fish.
CC Nucleic acids encoding the antimicrobial peptides are useful for treating
CC stress in a fish. Sequences AB07720-23 represent specific examples of
CC the antimicrobial peptides.
CC
SQ Sequence 22 AA:

Query Match 100.0%; Score 118; DB 23; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.8e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FHHIFRGIVHAGRSIGRFLTG 22
Db 1 FHHIFRGIVHAGRSIGRFLTG 22
|||||
RESULT 3
AAE20173 ID AAE20173 standard; peptide; 22 AA.
XX
AC AAE20173;
XX
DT 18-JUN-2002 (first entry)
XX
DE Fish mast cell antimicrobial peptide, piscidin 1 (P1).
XX
XX
KW Antimicrobial peptide; endobiotic peptide; mast cell; piscidin; fish;
KW methicillin; vancomycin; streptogramin; microbial infection; stress;
KW bacterial disease; fish health; screening; gene therapy; antimicrobial;
KW tranquilizer; fungicide; protozoicide; antibacterial; antiparasitic.
XX
OS Morone chrysops.
XX
OS Morone saxatilis.
XX
PN WO200214345-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001MO-US41696.
XX
PR 15-AUG-2000; 2000US-225354P.
XX
XX (UYNC-) UNTV NORTH CAROLINA STATE.
XX
PA Noga EJ, Silphaduang U;
XX
PI WPI; 2002-269176/31.
XX
XX
XX Novel antimicrobial peptides (endobiotic peptides) isolated from mast
XX PT cells, specifically fish mast cells, useful for treating microbial
XX PT infections in humans or animals and for reducing antibiotic resistance
XX PT in bacteria -
XX
PS Claim 4; Page 17; 31pp; English.
XX
XX The invention relates to antimicrobial peptide (endobiotic peptide)
XX CC isolated from a mast cell. Such peptides obtained from fish mast cells
XX CC are referred as piscidins. The antimicrobial peptide is useful for
XX CC treating microbial infection in a subject, and for reducing antibiotic
XX CC (e.g., methicillin, vancomycin or streptogramin) resistance in bacteria
XX CC such as *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus aureus*,
XX CC *Klebsiella pneumoniae*, *Streptococcus faecalis* or *Shigella flexneri*. These
XX CC peptides are useful for treating microbial infections in human or animals
XX CC and for preparing a medicament for treating microbial infections. These
XX CC peptides are also useful for treating infections caused by microorganisms
XX CC other than bacteria, e.g., fungi, yeast, protozoa, parasite, etc. These
XX CC peptides are useful for treating stress or bacterial disease in fish and
XX CC monitoring fish health. The inverse relationship between endobiotic
XX CC levels and stress also allows assessment of freshness of fish food

CC product. Nucleic acid sequences encoding these peptides are useful for
CC diagnostic purposes, treating stress in a fish and for screening
CC additional endobiotic peptides. They are also useful in gene therapy.
CC The present sequence is piscidin 1 (P1) peptide isolated from the gills
CC of hybrid striped bass (Morone saxatilis x Morone chrysops).

XX Sequence 22 AA;

Query Match 73.7%; Score 87; DB 23; Length 22;
Best Local Similarity 68.2%; Pred. No. 4.3e-07;
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 FTHHFRGIVHAGRSIGRFLTG 22
1 ||||| : : : : :
DB 1 FFHHFRGIVHVGKTIHRLVTG 22

RESULT 4
ABB07721
ID ABB07721 standard; peptide: 22 AA.

XX ABB07721;

DT 10-JUN-2002 (first entry)

DE Antimicrobial peptide #2 from hybrid striped bass.

XX Fish: hybrid striped bass; endobiotic; antibiotic; tranquilizer;
KM antimicrobial; antibacterial; fish food product.

XX Morone saxatilis x Morone chrysops.

XX WO200214346-A2.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US41697.

XX 15-AUG-2000; 2000US-225354P.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Noga EJ, Silphaduang U;

XX WPI; 2002-269177/31.

XX Novel antimicrobial compound (endobiotic peptide isolated from fish),
PT useful for monitoring fish health and monitoring freshness of fish food
PT product and for screening compounds which treat stress in fish -

XX Claim 1; Page 16; 25pp; English.

XX The invention provides antimicrobial peptides (ABB07720-23) isolated from
CC hybrid striped bass. The peptides were purified from the endobiotic
CC family. The endogenous endobiotic peptides exhibiting antimicrobial
CC activity) are useful for monitoring fish health, monitoring the freshness
CC of the fish product such as fresh, refrigerated, or frozen, fish food
CC product, and for screening compounds useful for treating stress in fish.
CC Nucleic acids encoding the antimicrobial peptides are useful for treating
CC stress in a fish. Sequences ABB07720-23 represent specific examples of
CC the antimicrobial peptides.

XX Sequence 22 AA;

Query Match 73.7%; Score 87; DB 23; Length 22;
Best Local Similarity 68.2%; Pred. No. 4.3e-07;
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 FTHHFRGIVHAGRSIGRFLTG 22
1 ||||| : : : : :
DB 1 FFHHFRGIVHVGKTIHRLVTG 22

RESULT 5
AAE20174
ID AAE20174 standard; peptide: 22 AA.

XX AAE20174;

XX 18-JUN-2002 (first entry)

DE Fish mast cell antimicrobial peptide, piscidin 2 (P2).

XX Antimicrobial peptide; endobiotic peptide; mast cell; piscidin; fish;
KM methicillin; vancomycin; streptogramin; microbial infection; stress;
KW bacterial disease; fish health; screening; gene therapy; antimicrobial;
KM tranquilizer; fungicide; protozoacide; antibacterial; antiparasitic.

XX Morone chrysops.
OS Morone saxatilis.

XX WO200214345-A2.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US41696.

XX 15-AUG-2000; 2000US-225354P.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Noga EJ, Silphaduang U;

XX WPI; 2002-269176/31.

XX Novel antimicrobial peptides (endobiotic peptides) isolated from mast
PT cells, specifically fish mast cells, useful for treating microbial
PT infections in humans or animals and for reducing antibiotic resistance
PT in bacteria -

XX Claim 4; Page 18; 31pp; English.

XX The invention relates to antimicrobial peptide (endobiotic peptide)
CC isolated from a mast cell. Such peptides obtained from fish mast cells
CC are referred as piscidins. The antimicrobial peptide is useful for
CC treating microbial infection in a subject, and for reducing antibiotic
CC (e.g., methicillin, vancomycin or streptogramin) resistance in bacteria
CC such as *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus aureus*,
CC *Klebsiella pneumoniae*, *Streptococcus faecalis* or *Shigella flexneri*. These
CC peptides are useful for treating microbial infections in human or animals
CC and for preparing a medicament for treating microbial infections. These
CC peptides are also useful for treating infections caused by microorganisms
CC other than bacteria, e.g., fungi, yeast, protozoa, parasite, etc. These
CC peptides are useful for treating stress or bacterial disease in fish and
CC monitoring fish health. The inverse relationship between endobiotic
CC levels and stress also allows assessment of freshness of fish food
CC product. Nucleic acid sequences encoding these peptides are useful for
CC diagnostic purposes, treating stress in a fish and for screening
CC additional endobiotic peptides. They are also useful in gene therapy.
CC The present sequence is piscidin 2 (P2) peptide isolated from the gills
CC of hybrid striped bass (Morone saxatilis x Morone chrysops).

XX Sequence 22 AA;

Query Match 71.2%; Score 84; DB 23; Length 22;
Best Local Similarity 63.6%; Pred. No. 1.3e-06;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 FTHHFRGIVHAGRSIGRFLTG 22
1 ||||| : : : : :
DB 1 FFHHFRGIVHVGKTIHRLVTG 22

RESULT 6
ABB07722
ID ABB07722 standard; peptide: 22 AA.

```

XX AC ABB07722;
XX XX 10-JUN-2002 (first entry)
XX DT
XX DE Antimicrobial peptide #3 from hybrid striped bass.
XX XX
XX DE Antimicrobial peptide #3 from hybrid striped bass.
XX XX
XX KW Fish: hybrid striped bass; endobiotic; antibacterial; tranquilizer;
XX KW antimicrobial; antibacterial; fish food product.
XX OS
XX PN WO200214346-A2.
XX PD 21-FEB-2002.
XX PF 13-AUG-2001; 2001WO-US41697.
XX PR 15-AUG-2000; 2000US-225354P.
XX PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX PI Noga EJ, Silphaduang U;
XX DR WPI; 2002-269177/31.
XX XX
XX PT Novel antimicrobial compound (endobiotic peptide isolated from fish),
XX PT useful for monitoring fish health and monitoring freshness of fish food
XX PT product and for screening compounds which treat stress in fish -
XX PS Claim 1; Page 16; 25pp; English.
XX CC
XX CC The invention provides antimicrobial peptides (ABB07720-23) isolated from
XX CC hybrid striped bass. The peptides were purified from the endobiotic
XX CC family. The endogenous endobiotic peptides exhibiting antimicrobial
XX CC activity are useful for monitoring fish health, monitoring the freshness
XX CC of the fish product such as fresh, refrigerated, or frozen, fish food
XX CC product, and for screening compounds useful for treating stress in fish.
XX CC Nucleic acids encoding the antimicrobial peptides are useful for treating
XX CC stress in a fish. Sequences ABB07720-23 represent specific examples of
XX CC the antimicrobial peptides.
XX SQ
XX Sequence 22 AA;
XX
XX Query Match 71.2%; Score 84; DB 23; Length 22;
XX Best Local Similarity 63.6%; Pred. No. 1.3e-06;
XX Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 FTHHIFRGIVHAGRSIGRFLTG 22
XX 1 ||||||||| : : |||
XX Db 1 FFHHIFRGIVHVKTIHKLVTG 22
XX
XX RESULT 7
XX AAE20177
XX ID AAE20177 standard; peptide; 24 AA.
XX XX
XX AC AAE20177;
XX XX
XX DT 18-JUN-2002 (first entry)
XX XX
XX DE Antimicrobial peptide analogue.
XX DE
XX KW Antimicrobial peptide; endobiotic peptide; mast cell; piscidin;
XX KW methicillin; vancomycin; streptogramin; microbial infection; stress;
XX KW bacterial disease; fish health; screening; gene therapy; antimicrobial;
XX KW tranquilizer; fungicide; protozoacide; antibacterial; antiparasitic.
XX OS
XX OS Unidentified.
XX XX
XX FH Key Location/Qualifiers
XX FT Disulfide-bond 12..13
XX XX

```

```

EN PN WO200214345-A2.
XX XX
XX PD 21-FEB-2002.
XX XX
XX PF 13-AUG-2001; 2001WO-US41696.
XX XX
XX PR 15-AUG-2000; 2000US-225354P.
XX XX
XX PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX XX
XX PI Noga EJ, Silphaduang U;
XX DR WPI; 2002-269176/31.
XX XX
XX PT Novel antimicrobial peptides (endobiotic peptides) isolated from mast
XX PT cells, specifically fish mast cells, useful for treating microbial
XX PT infections in humans or animals and for reducing antibiotic resistance
XX PT in bacteria -
XX PS
XX PS Example 9; Page 24; 31pp; English.
XX CC
XX CC The invention relates to antimicrobial peptide (endobiotic peptide)
XX CC isolated from a mast cell. Such peptides obtained from fish mast cells
XX CC are referred as piscidins. The antimicrobial peptide is useful for
XX CC treating microbial infection in a subject, and for reducing antibiotic
XX CC (e.g., methicillin, vancomycin or streptogramin) resistance in bacteria
XX CC such as Pseudomonas aeruginosa, Escherichia coli, Staphylococcus aureus,
XX CC Klebsiella pneumoniae, Streptococcus faecalis or Shigella flexneri. These
XX CC peptides are useful for treating microbial infections in human or animals
XX CC and for preparing a medicament for treating microbial infections. These
XX CC peptides are also useful for treating infections caused by microorganisms
XX CC other than bacteria, e.g., fungi, yeast, protozoa, parasite, etc. These
XX CC peptides are useful for treating stress or bacterial disease in fish and
XX CC monitoring fish health. The inverse relationship between endobiotic
XX CC levels and stress also allows assessment of freshness of fish food
XX CC product. Nucleic acid sequences encoding these peptides are useful for
XX CC diagnostic purposes, treating stress in a fish and for screening
XX CC additional endobiotic peptides. They are also useful in gene therapy.
XX CC The present sequence is antimicrobial peptide analogue.
XX SQ
XX Sequence 24 AA;
XX
XX Query Match 50.0%; Score 59; DB 23; Length 24;
XX Best Local Similarity 90.9%; Pred. No. 0.014;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 FTHHIFRGIVH 11
XX 1 |||||||||
XX Db 14 FFHHIFRGIVH 24
XX
XX RESULT 8
XX ABG25467
XX ID ABG25467 standard; Protein; 309 AA.
XX XX
XX AC ABG25467;
XX XX
XX DT 18-FEB-2002 (first entry)
XX XX
XX DE Novel human diagnostic protein #25458.
XX DE
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX XX
XX OS Homo sapiens.
XX OS
XX PN WO200175067-A2.
XX XX
XX PD 11-OCT-2001.
XX XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX XX
XX PR 31-MAR-2000; 2000US-0540217.
XX XX

```

PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 DR N-PSDB; AAS88564.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 55826; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in treating
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence . 309 AA:
 Query Match 41.5%; Score 49; DB 22; Length 309;
 Best Local Similarity 33.3%; Pred. No. 9.7;
 Matches 11; Conservative 6; Mismatches 4; Indels 12; Gaps 2;
 QY 1 FTIHIFR-----GIYHA-----GRSIGRFL 21
 DB 265 FVHHILRRNGIADGFRRLAALTIHGKAVGQYLF 297
 RESULT 9
 ABG24377
 ID ABG24377 standard; Protein: 454 AA.
 XX
 AC ABG24377;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #24368.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX

PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 DR N-PSDB; AAS88564.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 54736; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 454 AA:
 Query Match 41.5%; Score 49; DB 22; Length 454;
 Best Local Similarity 42.3%; Pred. No. 15;
 Matches 11; Conservative 2; Mismatches 5; Indels 8; Gaps 1;
 QY 3 HHIFRG-----TVAGRSIGRFL 20
 DB 426 HHRFRAQAQRVECVFRAGKEIGHFL 451
 RESULT 10
 AAU38243
 ID AAU38243 standard; Protein: 509 AA.
 XX
 AC AAU38243;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Salmomella typhi cellular proliferation protein #134.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Salmomella typhi.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 XX

PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS56102.

XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX

PS Example 3; Seq ID No 13836; 51pp; English.

XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequences are also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 509 AA;

Query Match 40.7%; Score 48; DB 22; Length 509;
 Best Local Similarity 52.9%; Pred. No. 25;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 6 FRGIYHAGRSIGRPLTG 22
 : | | | | | | | | | |
 Db 139 YRAIVEAHMEGRRFTG 155

RESULT 11
 AAM69358
 ID AAM69358 standard; Protein: 510 AA.
 XX
 AC AAM69358;

XX
 DT 01-DEC-1998 (first entry)

XX
 DE Pyridine nucleotide transhydrogenase subunit A.

KW 1,2-propanediol production; intracellular methylglyoxal conversion;
 KW unsaturated polyester resin; liquid laundry detergent; antifreeze;
 KW de-icing composition; animal feed; pyridine nucleotide transhydrogenase.
 OS
 XX Unidentified.

PN NO9837204-A1.

XX 27-AUG-1998.

XX 19-FEB-1998; 98MO-US03271.

XX 19-FEB-1997; 97US-0801344.

PA (WISC) WISCONSIN ALUMNI RES FOUND.
 XX

PI Altaras NE, Cameron DC, Shaw AJ;
 XX
 DR WPI; 1998-480799/41.
 DR N-PSDB; AAV58416.
 XX

XX
 PT Fermentative production of 1,2-propanediol from simple sugars -
 PT using recombinant microorganism that expresses enzymes that convert
 PT extracellular methylglyoxal, also new transformants and synthetic
 PT operons
 XX

PS Disclosure; Page 37-39; 59pp; English.

XX
 CC This sequence is the pyridine nucleotide transhydrogenase subunit A. The
 CC DNA can be used in a vector used in the method of the invention. The
 CC method is for the production of 1,2-propanediol (1), and comprises
 CC culturing a recombinant microorganism (A) that expresses enzymes(s) able
 CC to convert intracellular methylglyoxal (MG) to (1), in a medium
 CC containing a sugar, other than a 6-deoxyhexose, as carbon source. (1) is
 CC used in manufacture of unsaturated polyester resins, liquid laundry
 CC detergents, pharmaceuticals, cosmetics, antifreeze and de-icing
 CC compositions. By-products of the process (carbon dioxide, alcohols and
 CC organic acids) may be used e.g. as animal feeds. The method uses
 CC inexpensive sugars and provides production of (1) from renewable sources,
 CC without: (a) use of propylene oxide as intermediate; (b) generation of
 CC toxic waste; and (c) use of high temperature and pressure. Yields of (1)
 CC of up to 100 g/l can be achieved.
 XX

SQ Sequence 510 AA;

Query Match 40.7%; Score 48; DB 19; Length 510;
 Best Local Similarity 52.9%; Pred. No. 25;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 6 FRGIYHAGRSIGRPLTG 22
 : | | | | | | | | | |
 Db 139 YRAIVEAHMEGRRFTG 155

RESULT 12
 AAU34573
 ID AAU34573 standard; Protein: 510 AA.
 XX
 AC AAU34573;

XX
 DT 14-FEB-2002 (first entry)

XX
 DE E. coli cellular proliferation protein #154.

KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS *Escherichia coli*.

PN WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/70.

DR N-PSDB: AAS52432.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 10166; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 510 AA:

Query Match 40.7%; Score 48; DB 22; Length 510;
Best Local Similarity 52.9%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 6 FRGIYHAGRSIGRELTG 22
DB 139 TRALVEAHAEGRERFTG 155

RESULT 13
ABB54148
ID ABB54148 standard; Protein: 883 AA.
XX
AC ABB54148;
XX
DT 16-MAY-2002 (first entry)
XX
DE Lactococcus lactis protein yieh.
XX
KM Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
OS Lactococcus lactis IL1403.
XX
PN FR2807446-A1.
XX
PD 12-OCT-2001.
XX
PF 11-APR-2000; 2000FR-0004630.
XX
PR 11-APR-2000; 2000FR-0004630.
XX
PA (INRG) INRA INST MAT RECH AGRONOMIQUE.
XX
PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
DR WPI; 2002-043418/06.
XX
PT New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species -
XX
PS Claim 6; SEQ ID NO 850; 2504pp; French.
XX
CC The present invention is related to a Lactococcus lactis nucleotide
CC sequence (AAB590521) and related proteins (AAB53300-AAB55621). The

CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify *Lactococcus lactis* or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 883 AA:

Query Match 40.7%; Score 48; DB 23; Length 883;
Best Local Similarity 45.0%; Pred. No. 45;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 IHIFGIYHAGRSIGRELT 21
DB 51 IHSLYRNLIHSGSGSGLYLT 70

RESULT 14
AAV74582
ID AAV74582 standard; Protein: 512 AA.
XX
AC AAV74582;
XX
DT 21-MAR-2000 (first entry)
XX
DE *Neisseria gonorrhoeae* ORF 155 protein sequence SEQ ID NO:638.
XX
KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS *Neisseria gonorrhoeae*.
XX
PN WO957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Frasier C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappunli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
XX
DR N-PSDB: AAZ53344.
XX
PT Novel *Neisseria* polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 2; Page 444; 1453pp; English.
XX
AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAV74253 to AAV75941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as

Query Match	39.8%;	Score 4/;	DB 21;	Length 227
Best Local Similarity	41.28;	Pred. No. 36;		
Mismatches	4;	Indels	6;	Gaps 0;

6 FRGIVHAGRSIGF LTG 22
QY : : : : : :

Db 139 YRAVIEANAFGRFTG 155

AAV/4583	
ID	AAV74583 standard; Protein; 513 AA.

XX
AC
AAY74583;

21-MAR-2000 (first entry)

XX	Neisseria meningitidis ORF 155 protein sequence	SEQ ID NO:640
DE		

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; infection; meningitis; septicaemia
antigen; diagnosis; immunogenic; infection; meningitis; septicaemia

antibacterial; gene therapy.

Neisseria meningitidis.

AA
PN
XX
W09957280-A2.

11-NOV-1999
PD
XX

30-APR-1999; 99WO-US093346.

PR	01-MAY-1998;	98US-0083758
PR	31-JUL-1998;	98US-0094869

PR	02-SEP-1998;	98US-0098994
PR	02-SEP-1998;	98US-0099062
PR	02-SEP-1998;	98US-0100074

PR	09-OCT-1998;	98US-0103/49
PR	09-OCT-1998;	98US-0103794
		00US-0103796

PR	09-OCT-1998;	98DS-0105750
PR	25-FEB-1999;	99US-0121528

XX
PA (CHIR) CHIRON CORP.
INVENTOR : VINCE GENOMIC PES

PA (GENO-) INST GENOMIC RES:
XX

PI	Fraser C,	Ratti G,	Scalato E,	Scarselli M
PI	Petersen J,	Pizza M,	Rappuoli R,	
PI	Matteoli N,	Venter JC,		

PI Telephone #, Vendor ref.
XX
END: 2000-062150/05.
BT

DR N-PSDB; AAZ53345.

Novel Neisserial polypeptides predicted to be useful adjuvants for vaccines and diagnostics

XX
PS
Claim 2; Page 445; 1453pp; English

AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941

CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* fimbriae AA254537 and AA254576 and AA254616 to AA255473 represent CC and polypeptides. AA254537 and AA254576 are the present invention. The

CC PCR primers used in the exemplification of the present invention;
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the present invention as diagnostic reagents, and

CC the invention can be used as vaccines, as adjuvants or for treating or preventing infection due to immunogenic compositions. The polypeptides can be used in the form of a conjugate.

CC

CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.

XX
XX
XX Sequence 513 AA:

Query Match 39.8%; Score 47; DB 21; Length 5137

Best Local Similarity	41.28;	Pred. No. 36;	
Matches	7;	Conservative	4;
		Mismatches	6;
		Indels	0;
		Gaps	0

ZY 6 FRGI VHAGRSIGRELTG 22
Z : : : :
ZY

Db 139 YRAVIEAANAFGRFTG 15

3000

Search completed: May 11, 2003, 11:25:00
Job time : 78 secs

GenCore version 5.1.4-P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:23:54 ; Search time 28 Seconds
(without alignments)
23.118 Million cell updates/sec

Title: US-09-929-788-1
Perfect score: 118
Sequence: 1 FHHHFRGIVHAGRGTRRLG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	40.7	510	US-08-801-344-8	Sequence 8, Appl
2	48	40.7	510	US-09-498-599-8	Sequence 8, Appl
3	44	37.3	282	US-09-134-001C-3155	Sequence 3155, Ap
4	44	37.3	399	US-08-834-655-4	Sequence 4, Appl
5	44	37.3	399	US-08-834-033A-4	Sequence 4, Appl
6	44	37.3	399	US-09-363-526-4	Sequence 4, Appl
7	44	37.3	399	US-09-363-526-4	Sequence 20, Appl
8	44	37.3	399	US-09-350-235-20	Sequence 20, Appl
9	43	36.4	73	US-09-227-357-586	Sequence 356, App
10	43	36.4	471	US-09-134-001C-4904	Sequence 4904, Ap
11	42	35.6	362	US-08-513-974B-374	Sequence 374, App
12	41	34.7	885	US-08-042-747A-8	Sequence 8, Appl
13	41	34.7	885	US-08-804-439A-23	Sequence 23, Appl
14	41	34.7	885	US-08-720-229-23	Sequence 23, Appl
15	41	34.7	891	US-08-042-747A-6	Sequence 6, Appl
16	41	34.7	2296	US-08-286-819A-27	Sequence 27, Appl
17	41	34.7	2296	US-08-980-357-27	Sequence 27, Appl
18	40.5	34.3	170	US-09-130-663-2	Sequence 2, Appl
19	40.5	34.3	170	US-09-432-335-2	Sequence 2, Appl
20	40.5	34.3	170	US-09-614-022-2	Sequence 2, Appl
21	40	33.9	405	US-09-134-001C-4999	Sequence 4999, Ap
22	40	33.9	690	US-09-134-001C-4568	Sequence 4568, Ap
23	39.5	33.5	354	US-09-134-001C-3374	Sequence 3374, Ap
24	39	33.1	158	US-09-443-184-59	Sequence 59, Appl
25	39	33.1	187	US-08-713-825-1	Sequence 1, Appl
26	39	33.1	187	US-09-199-842-1	Sequence 1, Appl
27	39	33.1	327	US-07-867-105B-1	Sequence 1, Appl

28	39	33.1	376	US-09-056-556-202	Sequence 202, App
29	39	33.1	376	US-09-072-596-197	Sequence 197, App
30	39	33.1	567	US-08-504-459-2	Sequence 2, Appl
31	38	32.2	125	US-08-249-013-7	Sequence 7, Appl
32	38	32.2	125	US-08-886-863-7	Sequence 7, Appl
33	38	32.2	125	US-09-175-229-7	Sequence 7, Appl
34	38	32.2	125	US-09-175-229-7	Sequence 7, Appl
35	38	32.2	125	PCT-US95-06764-7	Sequence 7, Appl
36	38	32.2	159	US-09-280-839-3	Sequence 3, Appl
37	38	32.2	166	US-09-081-180-4	Sequence 4, Appl
38	38	32.2	166	US-09-040-786-4	Sequence 4, Appl
39	38	32.2	436	US-08-674-168-6	Sequence 6, Appl
40	38	32.2	622	US-09-311-626B-4	Sequence 4, Appl
41	37	31.4	865	US-07-803-633A-13	Sequence 13, Appl
42	37	31.4	129	US-08-513-974B-327	Sequence 327, App
43	37	31.4	357	US-09-149-476-755	Sequence 755, App
44	37	31.4	411	US-08-887-534A-80	Sequence 80, Appl
45	37	31.4	423	US-09-134-001C-5210	Sequence 5210, Ap
			436	US-08-146-010A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-801-344-8
Sequence 8, Application US/08801344
Patent No. 6087140
GENERAL INFORMATION:
APPLICANT: Cameron, Douglas C.
APPLICANT: Shaw, Anita J.
TITLE OF INVENTION: Microbial Production of 1,2-Propanediol
TITLE OF INVENTION: FROM SUGAR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Dewitt Ross & Stevens S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WI
COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,344
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REFERENCE/DOCKET NUMBER: 09820.037
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-831-2100
TELEFAX: 608-831-2106
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Pyridine nucleotide transhydrogenase, subunit
ORGANISM: A
US-08-801-344-8
Query Match 40.7%; Score 48; DB 3; Length 510;
Best Local Similarity 52.9%; Pred. No. 7.1;

Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 6 FRGIVHAGRSIGRLTG 22
: | | | | | | | | | |
Db 139 YRAIVEAHAEFGFRFTG 155

RESULT 2

US-09-498-599-8
; Sequence 8, Application US/09498599
; Patent No. 6303352
; GENERAL INFORMATION:

APPLICANT: Cameron, Douglas C.
APPLICANT: Shaw, Anita J.
APPLICANT: Altaras, Nedim E.

TITLE OF INVENTION: MICROBIAL PRODUCTION OF
TITLE OF INVENTION: 1,2-PROPANEDIOL FROM SUGAR
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:
ADDRESSEE: Demitt Ross & Stevens S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WI

COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,599
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REFERENCE/DOCKET NUMBER: 09820.037
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-831-2100
TELEFAX: 608-831-2106

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single

MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal

ORGANISM: Pyridine nucleotide transhydrogenase, subunit
ORGANISM: A

US-09-498-599-8

Query Match 40.7%; Score 48; DB 4; Length 510;
Best Local Similarity 52.9%; Pred. No. 7.1;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 6 FRGIVHAGRSIGRLTG 22
: | | | | | | | | | |
Db 139 YRAIVEAHAEFGFRFTG 155

RESULT 3
US-09-134-001C-3155
; Sequence 3155, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stramm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FHHIFRGIVH 11

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FHHIFRGIVH 11

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3155
LENGTH: 282
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3155

Query Match 37.3%; Score 44; DB 4; Length 282;
Best Local Similarity 41.2%; Pred. No. 16;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 FHHIFRGIVHAGRSIG 17
: | | | | | | | | | |
Db 188 FLHVFEQYVAGTELG 204

RESULT 4
US-08-834-655-4
; Sequence 4, Application US/08834655
; Patent No. 5968809
; GENERAL INFORMATION:

APPLICANT: KNOTZON, DEBORAH
APPLICANT: MURKERT, PRADIP
APPLICANT: HUANG, YONG-SHENG
APPLICANT: THURMUND, JENNIFER
APPLICANT: CHAUDHARY, SUNITA

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: RAE-VENTER LAW GROUP, P.C.
STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039
CITY: PALO ALTO
STATE: CA

COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,655
FILING DATE: 11-APR-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: RAE-VENTER, BARBARA
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: CGNE.124.00US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 328-4400
TELEFAX: (650) 328-4477
TELEX: N/A

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-834-655-4

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FHHIFRGIVH 11

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FHHIFRGIVH 11

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FHHIFRGIVH 11

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FHHIFRGIVH 11

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FHHIFRGIVH 11

Query Match 37.3%; Score 44; DB 2; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DB 326 FLDHMFHGIHV 336

1: 1:1 1111

RESULT 5
US-08-834-033A-4
Sequence 4, Application US/08834033A
Patent No. 6075183
GENERAL INFORMATION:
APPLICANT: KNOTZON, DEBORAH
APPLICANT: MUKERJI, PRADIP
APPLICANT: HUANG, YUNG-SHENG
APPLICANT: THURMOND, JENNIFER
APPLICANT: CHAUDHARY, SUNITA
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.
STREET: 2001 FERRY BUILDING
CITY: SAN FRANCISCO
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,033A
FILING DATE: 11-APR-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: WARD, MICHAEL R.
REGISTRATION NUMBER: 38,651
REFERENCE/DOCKET NUMBER: CGAB-300, USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 433-4150
TELEFAX: (415) 433-8716
TELEX: N/A
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-834-033A-4

Query Match 37.3%; Score 44; DB 3; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY 1 FIIHIFRGIHV 11
1: 1:1 1111
DB 326 FLDHMFHGIHV 336

RESULT 6
US-09-363-574-4
Sequence 4, Application US/09363574
Patent No. 6136574
GENERAL INFORMATION:
APPLICANT: KNOTZON, DEBORAH
APPLICANT: MUKERJI, PRADIP
APPLICANT: HUANG, YUNG-SHENG
APPLICANT: THURMOND, JENNIFER
APPLICANT: CHAUDHARY, SUNITA
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:

ADDRESSEE: LIMBACH AND LIMBACH L.L.P.
STREET: 2001 FERRY BUILDING
CITY: SAN FRANCISCO
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/363,574
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WARD, MICHAEL R.
REGISTRATION NUMBER: 38,651
REFERENCE/DOCKET NUMBER: CGAB-202 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 433-4150
TELEFAX: (415) 433-8716
TELEX: N/A
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-363-574-4

Query Match 37.3%; Score 44; DB 4; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY 1 FIIHIFRGIHV 11
1: 1:1 1111
DB 326 FLDHMFHGIHV 336

RESULT 7
US-09-363-526-4
Sequence 4, Application US/09363526
Patent No. 6410288
GENERAL INFORMATION:
APPLICANT: KNOTZON, DEBORAH
APPLICANT: MUKERJI, PRADIP
APPLICANT: HUANG, YUNG-SHENG
APPLICANT: THURMOND, JENNIFER
APPLICANT: CHAUDHARY, SUNITA
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: LIMBACH AND LIMBACH L.L.P.
STREET: 2001 FERRY BUILDING
CITY: SAN FRANCISCO
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/363,526
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WARD, MICHAEL R.
REGISTRATION NUMBER: 38,651

REFERENCE/DOCKET NUMBER: CGAB-201 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 433-4150
TELEFAX: (415) 433-8716
TELEX: N/A
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-363-526-4

Query Match 37.3%; Score 44; DB 4; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 FIHIFRGIVH 11
DB 326 FLDHMFHGIVH 336

RESULT 8
US-09-330-235-20
Sequence 20, Application US/09330235
Patent No. 6459018
GENERAL INFORMATION:
APPLICANT: Knutson, Debbie
TITLE OF INVENTION: POLYUNSATURATED FATTY ACIDS IN PLANTS
FILE REFERENCE: MOCO.156.0005
CURRENT APPLICATION NUMBER: US/09/330,235
PRIOR FILING DATE: 1999-06-10
PRIORITY FILING DATE: 1998-06-12
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 20
LENGTH: 399
TYPE: PRT
ORGANISM: Mortierella alpina
US-09-330-235-20

Query Match 37.3%; Score 44; DB 4; Length 399;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 FIHIFRGIVH 11
DB 326 FLDHMFHGIVH 336

RESULT 9
US-09-227-357-586
Sequence 586, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
PRIOR FILING DATE: 1999-01-08
PRIORITY FILING DATE: 1998-07-07
NUMBER OF SEQ ID NOS: 60/051,926
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 586
LENGTH: 73
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: SITE
LOCATION: (66)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-227-357-586

Query Match 36.4%; Score 43; DB 4; Length 73;
Best Local Similarity 43.8%; Pred. No. 5.6;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 1 FIHIFRGIVHGRSI 16
DB 326 FLDHMFHGIVHGRSI 16

DB 55 FAHLLIANVHXGSI 70

RESULT 10

US-09-134-001C-4904

Sequence 4904, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 4904

LENGTH: 471

TYPE: PR1

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4904

Query Match

Best Local Similarity 80.4%;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 IFRGIVHAGR 14

DB 361 IYRGIVHGR 370

RESULT 11

US-08-513-974B-374

Sequence 374, Application US/08513974B

Patent No. 6114139

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji

APPLICANT: Hosoya, Masaki

APPLICANT: Fujii, Ryo

APPLICANT: Ohtaki, Tetsuya

APPLICANT: Fukusumi, Shoji

APPLICANT: Ohgi, Kazuhiro

TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,

TITLE OF INVENTION: PRODUCTION, AND USE THEREOF

NUMBER OF SEQUENCES: 380

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/513,974B

FILING DATE: 14-SEP-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP95/01599

FILING DATE: 10-AUG-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-093989

FILING DATE: 19-AUG-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-057186

FILING DATE: 16-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-007177

FILING DATE: 20-JAN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-326611

FILING DATE: 28-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-270017

FILING DATE: 02-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-236357

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-236356

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189274

FILING DATE: 11-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189273

FILING DATE: 11-AUG-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189272

FILING DATE: 11-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Resnick, David S.

REGISTRATION NUMBER: 34,235

REFERENCE/DOCKET NUMBER: 45753

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 374:

SEQUENCE CHARACTERISTICS:

LENGTH: 362 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-513-974B-374

Query Match

Best Local Similarity 46.7%;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 4 HIFRGIVHAGRSIGR 18

DB 137 HRYTGIVHPLKSIIGR 151

RESULT 12

US-08-042-747A-8

Sequence 8, Application US/08042747A

Patent No. 5487969

GENERAL INFORMATION:

APPLICANT: Eberle, Richard

APPLICANT: Black, Darla

APPLICANT: Schicariello, Franco

APPLICANT: Hilliard, Julia K.

TITLE OF INVENTION: Cloning and Amplification of Monkey B

TITLE OF INVENTION: Virus Genes

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cox & Smith Incorporated

STREET: 112 East Pecan Street, Suite 2000

CITY: San Antonio

STATE: Texas

COUNTRY: USA

ZIP: 78205

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/042,747A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haymond, W. Bradley
REGISTRATION NUMBER: 35186
REFERENCE/DOCKET NUMBER: S-0072.179
TELECOMMUNICATION INFORMATION:
TELEPHONE: 210-554-5500
TELEFAX: 210-226-8395
TELEX: 767609
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-042-747A-8

Query Match 34.7%; Score 41; DB 1; Length 885;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 2 IHHRFGIVHAGRSIGRFLTG 22
Db 720 LYSFEGGLGVGRAVKVVG 740

RESULT 13
US-08-804-439A-23
Sequence 23, Application US/08804439A
Patent No. 6015565
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Ste 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,439A
FILING DATE: February 21, 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09176/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-804-439A-23

Query Match 34.7%; Score 41; DB 3; Length 885;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 2 IHHRFGIVHAGRSIGRFLTG 22
Db 720 LYSFEGGLGVGRAVKVVG 740

RESULT 14
US-08-720-229-23
Sequence 23, Application US/08720229
Patent No. 6022542
GENERAL INFORMATION:
APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229
FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-720-229-23

Query Match 34.7%; Score 41; DB 3; Length 885;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 2 IHHRFGIVHAGRSIGRFLTG 22
Db 720 LYSFEGGLGVGRAVKVVG 740

RESULT 15
US-08-042-747A-6
Sequence 6, Application US/08042747A
Patent No. 5487969
GENERAL INFORMATION:
APPLICANT: Eberle, Richard
APPLICANT: Black, Darla
APPLICANT: Scnicariello, Franco
APPLICANT: Hilliard, Julia K.
TITLE OF INVENTION: Cloning and Amplification of Monkey B
Virus Genes

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cox & Smith Incorporated
STREET: 112 East Pecan Street, Suite 2000
CITY: San Antonio
STATE: Texas
COUNTRY: USA
ZIP: 78205
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/042,747A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haymond, W. Bradley
REGISTRATION NUMBER: 35186
REFERENCE/DOCKET NUMBER: S-0072.179
TELECOMMUNICATION INFORMATION:
TELEPHONE: 210-554-5500
TELEFAX: 210-226-8395
TELEX: 767609
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 891 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-042-747A-6

Query Match 34.7%; Score 41; DB 1; Length 891;
Best Local Similarity 28.6%; Pred. No. 1.6e+02;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 2 IHHFRGIYHAGRSIGRFLTG 22
DB 722 LYSFEGIGDVGRAVGVKVMG 742

Search completed: May 17, 2003, 14:28:37
Job time : 30 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:27:19 ; Search time 54 Seconds
(without alignments)
39.295 Million cell updates/sec

Title: US-09-929-788-1

Perfect score: 118
Sequence: 1 FHHHFRGIVHGRSIGRFLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications-AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	22	9	US-09-929-788-1
2	87	73.7	22	9	US-09-929-788-2
3	84	71.2	22	9	US-09-929-788-3
4	80	67.8	23	9	US-10-076-816-42
5	48	40.7	509	10	US-09-815-242-13836
6	48	40.7	510	10	US-09-815-242-10166
7	44.5	37.7	1737	9	US-09-808-602-83
8	44.5	37.7	1737	9	US-09-800-198-71
9	44.5	37.7	2724	9	US-09-808-602-13
10	44.5	37.7	2724	9	US-09-800-198-13
11	44.5	37.7	2733	9	US-09-808-602-8
12	44.5	37.7	2733	9	US-09-800-198-8
13	44.5	37.7	2802	9	US-09-808-602-81
14	44.5	37.3	2802	9	US-09-800-198-69
15	44	37.3	491	10	US-09-997-664-3
16	43	36.4	43	10	US-09-864-761-35887
17	43	36.4	73	9	US-09-983-802-586
18	43	36.4	1436	10	US-09-815-242-5566
19	43	36.4	1442	10	US-09-815-242-12321

20	42	35.6	362	9	US-10-092-135-3	Sequence 3, Appl1
21	42	35.6	362	9	US-10-092-135-4	Sequence 4, Appl1
22	42	35.6	362	9	US-09-779-679-28	Sequence 28, Appl1
23	41	34.7	169	9	US-09-746-783-80	Sequence 80, Appl1
24	41	34.7	247	9	US-10-028-072-514	Sequence 514, Appl
25	41	34.7	247	9	US-10-174-590-604	Sequence 604, Appl
26	41	34.7	247	9	US-10-176-758-604	Sequence 604, Appl
27	41	34.7	247	9	US-10-175-737-604	Sequence 604, Appl
28	41	34.7	247	9	US-10-121-049-514	Sequence 514, Appl
29	41	34.7	247	9	US-10-123-904-514	Sequence 514, Appl
30	41	34.7	247	9	US-10-140-470-514	Sequence 514, Appl
31	41	34.7	247	9	US-10-173-706-604	Sequence 604, Appl
32	41	34.7	247	9	US-10-175-738-604	Sequence 604, Appl
33	41	34.7	247	9	US-10-175-752-604	Sequence 604, Appl
34	41	34.7	247	9	US-10-176-482-604	Sequence 604, Appl
35	41	34.7	247	9	US-10-176-757-604	Sequence 604, Appl
36	41	34.7	247	9	US-10-176-913-604	Sequence 604, Appl
37	41	34.7	247	9	US-10-180-557-604	Sequence 604, Appl
38	41	34.7	247	9	US-10-180-557-604	Sequence 604, Appl
39	41	34.7	247	9	US-10-173-700-604	Sequence 604, Appl
40	41	34.7	247	9	US-10-174-572-604	Sequence 604, Appl
41	41	34.7	247	9	US-10-174-579-604	Sequence 604, Appl
42	41	34.7	247	9	US-10-174-582-604	Sequence 604, Appl
43	41	34.7	247	9	US-10-174-588-604	Sequence 604, Appl
44	41	34.7	247	9	US-10-175-739-604	Sequence 604, Appl
45	41	34.7	247	9	US-10-175-740-604	Sequence 604, Appl

ALIGNMENTS

RESULT 1
US-09-929-788-1
Sequence 1, Application US/09929788
Publication No. US20030083247A1
GENERAL INFORMATION:
APPLICANT: No. US20030083247A1a, Edward J.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES ISOLATED FROM FISH
FILE REFERENCE: 5051.519
CURRENT FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: US 60/225,354
PRIOR FILING DATE: 2000-08-15
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 22
TYPE: PRT
ORGANISM: Morone saxatilis x Morone chrysops
US-09-929-788-1

Query Match Score 118; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 8; 5e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FHHHFRGIVHGRSIGRFLTG 22
Db 1 FHHHFRGIVHGRSIGRFLTG 22
RESULT 2
US-09-929-788-2
Sequence 2, Application US/09929788
Publication No. US20030083247A1
GENERAL INFORMATION:
APPLICANT: No. US20030083247A1a, Edward J.
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES ISOLATED FROM FISH
FILE REFERENCE: 5051.519
CURRENT FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: US 60/225,354

```
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Morone saxatilis x Morone chrysops
US-09-929-788-2
```

```
Query Match
Best Local Similarity 67.8%; Score 87; DB 9; Length 22;
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 FHHHFRGIVHAGRSIGRFLTG 22
      | | | | | | | | | | | | | | | | | | | | | |
Db 1 FHHHFRGIVHAGKTIHDLVLTG 22
```

```
RESULT 3
US-09-929-788-3
; Sequence 3, Application US/09929788
; Publication No. US20030083247A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030083247A1, Edward J.
; APPLICANT: Silphaduang, Umaporn
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES ISOLATED FROM FISH
; FILE REFERENCE: 5051.519
; CURRENT APPLICATION NUMBER: US/09/929,788
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,354
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Morone saxatilis x Morone chrysops
US-09-929-788-3
```

```
Query Match
Best Local Similarity 71.2%; Score 84; DB 9; Length 22;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 FHHHFRGIVHAGRSIGRFLTG 22
      | | | | | | | | | | | | | | | | | | | | | |
Db 1 FHHHFRGIVHAGKTIHDLVLTG 22
```

```
RESULT 4
US-10-076-816-42
; Sequence 42, Application US/10076816
; Publication No. US20030056244A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ming
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076,816
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,188
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 23
; TYPE: PRT
```

```
; ORGANISM: Morone sp.
US-10-076-816-42
```

```
Query Match
Best Local Similarity 67.8%; Score 80; DB 9; Length 23;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 FHHHFRGIVHAGRSIGRFLTG 22
      | | | | | | | | | | | | | | | | | | | | | |
Db 1 FHHHFRGIVHAGKTIHDLVLTG 22
```

```
RESULT 5
US-09-815-242-13836
; Sequence 13836, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13836
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Salmonella typhl
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(509)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13836
```

```
Query Match
Best Local Similarity 40.7%; Score 48; DB 10; Length 509;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 6 FRCIVHAGRSIGRFLTG 22
      : | | | | | | | | | | | | | | | | | | | | | |
Db 139 YRAIVEAHEFGFRFTG 155
```

```
RESULT 6
US-09-815-242-10166
; Sequence 10166, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
```

```

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 10166
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10166
```

```

Query Match          40.7%; Score 48; DB 10; Length 510;
Best Local Similarity 52.9%; Pred. No. 32;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
```

```

Qy 6 FNGYVHAGSGTGRFLT 22
Db 139 YRAIVEAHEFGFPTG 155
```

```

RESULT 7
US-09-808-602-83
; Sequence 83, Application US/09808602
; Patent No. US2002015115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US2002015115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 83
; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-83
```

```

Query Match          37.7%; Score 44.5; DB 9; Length 1737;
Best Local Similarity 45.5%; Pred. No. 3.7e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;
```

```

Qy 3 HHIFR---GIVHAGSGTGRFLT 21
Db 147 HHILNVKSGILHKGTGENQFLT 168
```

```

RESULT 8
US-09-800-198-71
; Sequence 71, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 71
; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-198-71
```

```

Query Match          37.7%; Score 44.5; DB 9; Length 1737;
Best Local Similarity 45.5%; Pred. No. 3.7e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;
```

```

Qy 3 HHIFR---GIVHAGSGTGRFLT 21
Db 147 HHILNVKSGILHKGTGENQFLT 168
```

```

RESULT 9
US-09-808-602-13
; Sequence 13, Application US/09808602
; Patent No. US2002015115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US2002015115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-13
```

```

Query Match          37.7%; Score 44.5; DB 9; Length 2724;
Best Local Similarity 45.5%; Pred. No. 5.9e+02;
```

Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 3 HHFR---GIVHAGRSIGRFLT 21

Db 1134 HHILNVKSGILHKGTGENOFLLT 1155

RESULT 10

US-09-800-198-13

; Sequence 13, Application US/09800198

; Publication No. US20030087816A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie AM

; APPLICANT: Fernandes, Elma

; APPLICANT: Shinkets, Richard A

; APPLICANT: Herrmann, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishna

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-697

; CURRENT APPLICATION NUMBER: US/09/800,198

; CURRENT FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13

; LENGTH: 2724

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-800-198-13

Query Match 37.7%; Score 44.5; DB 9; Length 2724;

Best Local Similarity 45.5%; Pred. No. 5.9e+02;

Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 3 HHFR---GIVHAGRSIGRFLT 21

Db 1134 HHILNVKSGILHKGTGENOFLLT 1155

RESULT 11

US-09-808-602-8

; Sequence 8, Application US/09808602

; Patent No. US20020155115A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie A

; APPLICANT: Fernandes, Elma

; APPLICANT: Shinkets, Richard A

; APPLICANT: Herrman, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishnu

; APPLICANT: Mezes, Peter S

; APPLICANT: MacDougall, John

; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 15966-697 CIP

; CURRENT APPLICATION NUMBER: US/09/808,602

; CURRENT FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 09/800,198

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 2733

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-808-602-8

Query Match 37.7%; Score 44.5; DB 9; Length 2733;

Best Local Similarity 45.5%; Pred. No. 5.9e+02;

Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 3 HHFR---GIVHAGRSIGRFLT 21

Db 1143 HHILNVKSGILHKGTGENOFLLT 1164

RESULT 12

US-09-800-198-8

; Sequence 8, Application US/09800198

; Publication No. US20030087816A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie AM

; APPLICANT: Fernandes, Elma

; APPLICANT: Shinkets, Richard A

; APPLICANT: Herrmann, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishna

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-697

; CURRENT APPLICATION NUMBER: US/09/800,198

; CURRENT FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 2733

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-800-198-8

Query Match 37.7%; Score 44.5; DB 9; Length 2733;

Best Local Similarity 45.5%; Pred. No. 5.9e+02;

Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 3 HHFR---GIVHAGRSIGRFLT 21

Db 1143 HHILNVKSGILHKGTGENOFLLT 1164

RESULT 13

US-09-808-602-81

; Sequence 81, Application US/09808602

; Patent No. US20020155115A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie A

; APPLICANT: Fernandes, Elma

; APPLICANT: Shinkets, Richard A

; APPLICANT: Herrman, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishnu

; APPLICANT: Mezes, Peter S

; APPLICANT: MacDougall, John

; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 15966-697 CIP

; CURRENT APPLICATION NUMBER: US/09/808,602

; CURRENT FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 09/800,198

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 81

; LENGTH: 2802

; TYPE: PRT

; ORGANISM: Gallus gallus

US-09-808-602-81

Best Local Similarity 35.7%; Pred. No. 1.2e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Job time : 56 secs

ORGANISM: Gallus gallus
US-09-800-198-69

Query Match	37.78;	Score 44.5;	DB 9;	Length 2802;
Best Local Similarity	40.9%;	Pred. No. 6.1e+02;		
Matches	9;	Conservative	3;	Mismatches 7;
				Indels 3

Qy	3	HHFR--GIVHAGRSIGRFLT	21
		: : :	
Db	1212	HHVLNVKSGLHKNGENQFLT	1233

```

      RESULT 15
      US-09-997-664-3
      ; Sequence 3, Application US/09997664
      ; Patent No. US20020151003A1
      ; GENERAL INFORMATION:
      ; APPLICANT: Ben-Bassat, Arle
      ; APPLICANT: Cattermole, Monica
      ; APPLICANT: Gatenby, Anthony A.
      ; APPLICANT: Gibson, Katherine J.
      ; APPLICANT: Ramos-Gonzalez, Isabel
      ; APPLICANT: Ramos, Juan
      ; APPLICANT: Seriaslani, Sina
      ; TITLE OF INVENTION: Method for the Production of p-Hydroxybenzoate in Species of
      ; FILE REFERENCE: BC1018 US CIP
      ; CURRENT APPLICATION NUMBER: US/09/997,664
      ; CURRENT FILING DATE: 2001-11-28
      ; PRIOR APPLICATION NUMBER: 09/585,174
      ; PRIOR FILING DATE: 2000-06-01
      ; NUMBER OF SEQ ID NOS: 112
      ; SOFTWARE: Microsoft Office 97
      ; SEQ ID NO 3
      ; LENGTH: 491
      ; TYPE: PRF
      ; ORGANISM: Pseudomonas mendocina KR-1
      ; US-09-997-664-3

```

query match	37.3%;	Score 44;	DB 10;	Length 491;
-------------	--------	-----------	--------	-------------

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:22:54 ; Search time 43 Seconds
(without alignments)
49.185 Million cell updates/sec

Title: US-09-929-788-1

Perfect score: 118
Sequence: 1 FIIHIFRGIVHAGRSIGRFLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	44.9	524	2	H97658
2	53	44.9	524	2	AB2883
3	49	41.5	917	1	VGBEBH
4	48	40.7	83	2	AC2951
5	48	40.7	508	2	AH0280
6	48	40.7	509	2	AB0683
7	48	40.7	510	1	DEECXA
8	48	40.7	883	2	C86729
9	47.5	40.3	347	2	AG2840
10	47.5	40.3	357	2	AG7618
11	47	39.8	512	2	D82444
12	47	39.8	513	2	F81134
13	47	39.8	513	2	C81885
14	47	39.8	1763	2	T17465
15	46	39.0	360	2	T11067
16	46	39.0	530	2	S77433
17	46	39.0	687	2	D86314
18	46	39.0	880	2	AD1953
19	45.5	38.6	388	1	CBZM
20	45.5	38.6	392	1	CBVF
21	45.5	38.6	397	1	CBMZ
22	45	38.1	141	2	C82466
23	45	38.1	1041	2	T15521
24	45	38.1	3413	2	T17467
25	45	38.1	5069	2	T17464
26	44.5	37.7	412	2	H75484
27	44.5	37.7	417	2	AR3056
28	44.5	37.7	423	2	AR9230
29	44	37.3	271	2	G89929

ALIGNMENTS

RESULT 1

H97658

pyridine nucleotide transhydrogenase alpha chain (AY026033) [imported] - Agrobacterium

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: H97658

R:Goodner, B.; Hinkle, G.; Galtung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldm

A.: Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,

Science 284, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: H97658

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-524 <KUR>

A:Cross-references: GB:AB007869; PIDN:AAK88225.1; PID:g15157679; GSPDB:GN00169

A:Map position: circular chromosome

C:Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogena

Query Match 44.9%; Score 53; DB 2; Length 524;

Best Local Similarity 47.1%; Pred. No. 2.9;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

DB 141 YRAVTEAGNNNGRFTG 157

AB2883

NAD(P)+ transhydrogenase puta [imported] - Agrobacterium tumefaciens (strain C58, Dup

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AB2883

C:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyaev, T.; Levy, R.; Li, M.; McCl

: Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AB2883

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-524 <KUR>

A:Cross-references: GB:AB008688; PIDN:AAU43480.1; PID:g17740986; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: puta

hypothetical prote
syn protein - Rhi
tryptophanyl-tRNA
carcinoembryonic a
probable homobox
conserved hypothet
4-hydroxybenzoate
probable anion tra
hypothetical prote
protoporphyrin IX
hypothetical prote
hypothetical prote
conserved hypothet
ribosomal protein
probable aldehyde
p-hydroxybenzaldehy

C>Date: 30-Jun-1988 #sequence.revision 19-May-1995 #extl_change 03-Jun-2002
C:Accession: S24380; A25012; E64916; S18355; S23853
R:Almad, S.; Glavas, N.A.; Bragg, P.D.
Eur. J. Biochem. 207, 733-739, 1992
A>Title: A mutation at G1314 of the beta subunit of the Escherichia coli pyridine nucleotide transhydrogenase
A:Reference number: S24380; MUID:92339464; PMID:1633824
A:Accession: S24380
Molecule type: DNA
A:Residues: 1-510 <AAH>
A:Cross-references: EMBL:X66086; NID:g42455; PIDN:CAA46884.1; PID:g42456
R:Clarke, D.M.; Loo, T.W.; Gillam, S.; Bragg, P.D.
Eur. J. Biochem. 158, 647-653, 1986
A>Title: Nucleotide sequence of the pntA and pntB genes encoding the pyridine nucleotide transhydrogenase
A:Reference number: A91172; MUID:86274751; PMID:3525165
A:Accession: A25012
Molecule type: DNA
A:Residues: 1-39; VNMQVLTIRKLCG', 56-328, 'R', 330-502 <CLA>
R:Baltner, F.R.; Plunkett III, G.; Bloch, C.A.; Pernin, N.T.; Burland, V.; Riley, M.; Collier, R.C.W.; Glavas, N.A.; Bragg, P.D.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:9742661; PMID:9278503
A:Accession: E64916
A>Status: nucleic acid sequence not shown; translation not shown
Molecule type: DNA
A:Residues: 1-510 <BAAT>
A:Cross-references: GB:E000255; GB:U00096; NID:g1787875; PIDN:AAC74675.1; PID:g1787887
A:Experimental source: strain K-12, substrain MG1655
R:Tong, R.C.W.; Glavas, N.A.; Bragg, P.D.
Biochim. Biophys. Acta 1080, 19-28, 1991
A>Title: Topological analysis of the pyridine nucleotide transhydrogenase of Escherichia coli K-12
A:Reference number: S18355; MUID:92031646; PMID:1932078
A:Accession: S18355
Molecule type: protein
A:Residues: 1-10; 16-25; 'XF', 'ISV', 136-138, 225, 'EGT', 229-238; 270-273, 'X', 275-278; 'AG', 279-280
C:Genetics:
A:Gene: pntA
A:Map position: 35 min
C:Complex: heterotetramer; two alpha and two beta chains
C:Function:
A>Description: catalyzes the reversible hydride ion transfer between NAD and NADP; this reaction

A:Pathway: NAD phosphorylation and dephosphorylation
C:Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogenase
C:Keywords: heterotetramer; inner membrane; transmembrane protein; NAD; NADP; oxidoreductase
F:1-509/Domain: NAD(P)+ transhydrogenase (B-specific) alpha chain homology <TBA>
F:1-295/Domain: alanine dehydrogenase homology <ALA>
F:167-195/Region: beta-alpha-beta NAD nucleotide-binding fold
F:404-420/Domain: transmembrane #status predicted <TM1>
F:429-445/Domain: transmembrane #status predicted <TM2>
F:453-465/Domain: transmembrane #status predicted <TM3>
F:478-494/Domain: transmembrane #status predicted <TM4>

Query Match	Score 40.7%	DB 1:	Length 510:
Best Local Similarity 52.9%;	Pred. NO. 16;		
Matches 9; Conservative 1;	Mismatches 7;	Indels 0;	Gaps 0;
QY 6 FGCIIVAGRSIGRELTG 22			
: I I I I I I I I I I			
DB 139 YRAIVEAHAEFGREFTG 155			

RESULT 8
C86729
hypothetical protein yieh [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: C86729
R:Polcino, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, R.; Bolotin, A.; Wincker, P.
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: C86729

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-883 <SNO>
A:Cross-references: GB:AE005176; PID:g12723760; PIDN:AAK04933.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: y1eh

Query Match          40.7%; Score 48; DB 2; Length 883;
Best local similarity 45.0%; Pred. No. 29;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2  IHHRFGIVHAGRSIGRFLT 21
      ||::||::||::||::||
Db       51  IHSLYRNILHSGSSGFLYT 70

RESULT 9
AG2840
Alcohol dehydrogenase adh [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AG2840
R:Wood, D.W.; Seibubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McGill
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG2840
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-347 <K0X>
A:Cross-references: GB:AE008688; PIDN:AAU43141.1; PID:g17740616; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: adh
A:Map position: circular chromosome
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match          40.3%; Score 47.5; DB 2; Length 347;
Best local similarity 50.0%; Pred. No. 13;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY      4  HIRGIYVHA-GRSIGRFLT 22
      |::|::|::|::|::|
Db       60  HEFAGVISAGRNVTFRKGT 79

RESULT 10
A97618
Alcohol dehydrogenase (adh-hc) [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: A97618
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: A97618
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87898.1; PID:g15157292; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3897
A:Map position: circular chromosome
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match          40.3%; Score 47.5; DB 2; Length 357;

```

Best Local Similarity 50.0%; Pred. No. 13;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 4 HIFRGIIVHA-GSISGRFLTG 22
| | | | | | | | | | | | | | | | | | | | | |
Db 70 HEFAGVISAGKRVNFRFTG 89

RESULT 11

D82444
NAD(P) transhydrogenase, alpha chain VCA0563 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82444
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;
Charadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82444
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-512 <HEI>
A:Cross-references: GB:AE004387; GB:AE003853; NID:99657967; PIDN:AAF96465.1; GSPDB:GN001
C:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0563
A:Map position: 2
C:Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogenase

Query Match 39.8%; Score 47; DB 2; Length 512;
Best Local Similarity 47.1%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIIVHGRSISGRFLTG 22
| | | | | | | | | | | | | | | | | | | | | |
Db 139 YRAVIEAANAIFGRFTG 155

RESULT 12

F81134
NAD(P) transhydrogenase, alpha chain NMB0980 [imported] - Neisseria meningitidis (strain
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: F81134
R:Letellin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizsa, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; VE
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: F81134
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-513 <NET>
A:Cross-references: GB:AE002449; GB:AE002098; NID:97226218; PIDN:AAF41384.1; PID:9722622
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0980
C:Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogenase
F:1-512/Domain: NAD(P)+ transhydrogenase (B-specific) alpha chain homology <TBA>
F:1-296/Domain: alanine dehydrogenase homology <ALA>

Query Match 39.8%; Score 47; DB 2; Length 513;
Best Local Similarity 41.2%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 6 FRGIIVHGRSISGRFLTG 22
| | | | | | | | | | | | | | | | | | | | | |
Db 139 YRAVIEAANAIFGRFTG 155

RESULT 13

C81885
probable NAD(P) transhydrogenase (B-specific) (EC 1.6.1.1) alpha chain NMA1177 [imported]
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C:Accession: C81885
R:Fairchild, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: C81885
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-513 <PAR>
A:Cross-references: GB:AL162755; GB:AL157959; NID:97379742; PIDN:CA884439.1; PID:9737
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: pntA; NMA1177
C:Superfamily: NAD(P)+ transhydrogenase (B-specific) alpha chain; alanine dehydrogena
C:Keywords: oxidoreductase
F:1-512/Domain: NAD(P)+ transhydrogenase (B-specific) alpha chain homology <TBA>
F:1-296/Domain: alanine dehydrogenase homology <ALA>

Query Match 39.8%; Score 47; DB 2; Length 513;
Best Local Similarity 41.2%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 6 FRGIIVHGRSISGRFLTG 22
| | | | | | | | | | | | | | | | | | | | | |
Db 139 YRAVIEAANAIFGRFTG 155

RESULT 14

T17465
rifamycin polyketide synthase module 7 - Amycolatopsis mediterranei
C:Species: Amycolatopsis mediterranei
C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 03-Nov-2000
C:Accession: T17465
R:Schupp, T.
submitted to the EMBL Data Library, December 1997
A:Reference number: Z18802
A:Accession: T17465
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1763 <SCH>
A:Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227122; PIDN:CAA11037.1
A:Experimental source: strain LBG A3136
C:Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-ox
homology; [acyl-carrier-protein] S-malonyltransferase homology
C:Keywords: carrier protein
F:1615-1686/Domain: acyl carrier protein homology <ACP>

Query Match 39.8%; Score 47; DB 2; Length 1763;
Best Local Similarity 47.1%; Pred. No. 85;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 IHHIFRGIIVHGRSISGR 18
| | | | | | | | | | | | | | | | | | | | | |
Db 1576 VPHILRGLVIRGRQAR 1592

RESULT 15

T11067
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - nematode (Onchocerca vo
C:Species: Onchocerca volutus
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11067
R:Keddie, E.M.; Higazi, T.; Unnasch, T.R.
Mol. Biochem. Parasitol. 95, 111-127, 1998
A:Title: The mitochondrial genome of Onchocerca volutus: Sequence, structure and phy
A:Reference number: Z17246; MUID:98434257; PMID:9763293

A:Accession: T11067
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-360 <KED>
A:Cross-references: EMBL:AF015193; NID:g2735934; PID:g2735939; PIDN:AAC61614.1
A:Experimental source: strain forest
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:1-196/Domain: cytochrome b6 homology <CR6>
F:70-169/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:64-183/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 39.0%; Score 46; DB 2; Length 360;
Best local similarity 45.8%; Pred. No. 23;
Matches 11; Conservative 7; Mismatches 4; Indels 2; Gaps 2;

QY 1 FTH-HIFRGIVHAG-RSIGRFLTG 22
11:||||:|11:11:
Db 80 FYLHIFKGLIYGSYRLIGVWLSG 103

Search completed: May 17, 2003, 14:28:01
Job time : 45 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:15:39 ; Search time 25 Seconds
(without alignments)
36.499 Million cell updates/sec

Title: US-09-929-788-1
Perfect score: 118
Sequence: 1 FHHIFRGIHAGRSIGRFLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.5	41.5	917	1	VGLB_HSVB2
2	48.0	40.7	510	1	PMTA_ECOLI
3	45.5	38.6	388	1	CYB_MALIZ
4	45.5	38.6	392	1	CYB_SOLTU
5	45.5	38.6	392	1	CYB_YICFA
6	45.5	38.6	394	1	CYB_OENBE
7	45.5	38.6	397	1	CYB_ORYSA
8	45.5	38.1	666	1	PD14_MOUSE
9	44.4	37.3	336	1	STRM_RHLET
10	44.4	37.3	380	1	STRM_RHLET
11	44.4	37.3	776	1	PSAB_HETTR
12	43.3	36.4	395	1	ARP2_CAEEL
13	43.3	36.4	445	1	NSMA_CAEEL
14	43.3	36.4	492	1	GYD2_HAINT
15	43.3	36.4	492	1	GYD2_HAINT
16	43.3	36.4	663	1	PD15_HUMAN
17	43.3	36.4	1436	1	DPO3_STAM
18	43.3	36.4	1438	1	DPO3_STAM
19	42.5	36.0	515	1	CXAA_HUMAN
20	42.5	35.6	354	1	SYL_DROSO
21	42.5	35.6	362	1	P2YR_CHICK
22	42.5	35.6	362	1	P2YR_CHICK
23	42.5	35.6	378	1	P2YR_CHICK
24	42.5	35.6	666	1	H182_CAOCR
25	41.5	35.2	467	1	LEU2_VIRCH
26	41.5	34.7	114	1	KDGL_BACSU
27	41.5	34.7	201	1	GCH1_STRO
28	41.5	34.7	270	1	TRUA_ECOLI
29	41.5	34.7	411	1	EF1G_TYR
30	41.5	34.7	521	1	NU2C_SYNP
31	41.5	34.7	586	1	HO_YEAST
32	41.5	34.7	885	1	VGLB_HSV2S
33	40.5	34.3	170	1	OBPB_HUMAN

34	40.5	34.3	250	1	RK13_SPIOL
35	40	33.9	181	1	NDKM_COLLI
36	40	33.9	219	1	CLNI_HUMAN
37	40	33.9	219	1	CLNI_MOUSE
38	40	33.9	225	1	GYD2_YEAST
39	40	33.9	322	1	GDC_RAT
40	40	33.9	330	1	GDC_BOVIN
41	40	33.9	425	1	Y461_MYCGE
42	40	33.9	425	1	Y461_MYCPN
43	40	33.9	512	1	Y461_HAEIN
44	40	33.9	676	1	NTP2_VACCA
45	40	33.9	676	1	NTP2_VACCC

ALIGNMENTS

RESULT 1	ID	VGLB_HSVB2	STANDARD:	PRT:	917 AA.
AC	P12641:				
DT	01-OCT-1989 (Rel. 12, Created)				
DT	01-AUG-1990 (Rel. 15, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Glycoprotein B-1 precursor.				
OS	Bovine herpesvirus type 2 (strain BMV) (Bovine gammaherpesvirus 2)				
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae.				
OC	Alphaherpesvirinae; Simplexvirus.				
OX	NCBI_TaxID=10296;				
RN	[1]				
RX	SEQUENCE FROM N.A.				
RA	MEDLINE=8306231; PubMed=2841793;				
RA	Hammerhead W., Cornatns F., Mankertz J., Pauli G., Ludwig H.,				
RA	Buhk H.-J.,				
RT	"Conservation of a gene cluster including glycoprotein B in bovine				
RT	herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";				
RL	Virology 165:388-405(1988).				
RN	[2]				
RP	SEQUENCE OF 1-200 FROM N.A.				
RX	MEDLINE=88306232; PubMed=2457278;				
RA	Hammerhead W., Cornatns F., Mankertz J., Buhk H.-J., Pauli G., Ludwig H.,				
RA	"Common epitopes of glycoprotein B map within the major DNA-binding				
RT	proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex				
RT	virus type 1 (HSV-1).";				
RL	Virology 165:406-418(1988).				
CC	-1- FUNCTION: GRI IS A 130 kDa GLYCOPROTEIN WHICH IS NECESSARY FOR THE				
CC	PENETRATION OF THE VIRUS INTO THE HOST CELL AND THE INDUCTION OF A				
CC	SYNCYTIAL PHENOTYPE.				
CC	-1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: M21628; AAA46053.1; -				
DR	EMBL: M21632; AAA46052.1; -				
DR	PIR: C29242; VGBEBH.				
DR	InterPro: IPR000234; Glycoprot. B.				
DR	Pfam: PF00606; Glycoprotein.B; 1.				
DR	ProDom: PD000693; Glycoprot. B; 1.				
KW	Glycoprotein; Transmembrane; Signal.				
FT	SIGNAL	1	22		
FT	CHAIN	23	917		GLYCOPROTEIN B-1.
FT	TRANSMEM	578	594		POTENTIAL.
FT	TRANSMEM	770	786		POTENTIAL.
FT	TRANSMEM	795	811		POTENTIAL.
FT	CARBOHYD	48	48		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	110	110		N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	164	164	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	278	278	N-LINKED (GLCNAC. .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	421	421	N-LINKED (GLCNAC. .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	453	453	N-LINKED (GLCNAC. .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	505	505	N-LINKED (GLCNAC. .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	564	564	N-LINKED (GLCNAC. .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD	692	692	N-LINKED (GLCNAC. .) <th>(POTENTIAL).</th>	(POTENTIAL).
SO	SEQUENCE	917 AA;	101882 MW;	1B96CBF50DB4D3FC	CRC64;

Query Match	Best Local Similarity	41.5%;	Score 49;	DB 1;	Length 917;			
Matches	8;	Conservative	5;	Mismatches	8;	Indels	0;	Gaps
QY	2	IHHIFGCIYHAGRSIGRFLTG	22					
Db	753	LHGFEELGSDYGRVGRVVLG	773					


```

RESULT 2
PNTA_ECOLI
ID PNTA_ECOLI STANDARD: PRT: 510 AA.
AC P07001: P76888:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE MD(P) transhydrogenase subunit alpha (BC 1.6.1.2) (Pyridine
DE nucleotide transhydrogenase subunit alpha) (Nicotinamide nucleotide
DE transhydrogenase subunit alpha).
GN PNTA OR B1603.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562:
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86274751; PubMed=3525165;
RA Clarke D.M., Loo T.W., Gillam S., Bragg P.D.;
RT "Nucleotide sequence of the pntA and pntB genes encoding the pyridine
RT nucleotide transhydrogenase of Escherichia coli."
RL Eur. J. Biochem. 158:647-653(1986).
[2]
RN REVISIONS.
RP MEDLINE=92339464; PubMed=1633824;
RX Ahmad S., Glavas N.A., Bragg P.D.;
RT "A mutation at Gly314 of the beta subunit of the Escherichia coli
RT pyridine nucleotide transhydrogenase abolishes activity and affects
RT the NADP(H)-induced conformational change."
RL Eur. J. Biochem. 207:733-739(1992).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burlew V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasaai H., Kashimoto K., Kimura S., Kitagawa M.,
RA Kitagawa K., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakano S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubraman S.,
RA Tachina H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamanoto Y., Horiiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map."
RL DNA Res. 3:363-377(1996).
-1- FUNCTION: THE TRANSHYDROGENATION BETWEEN NADH AND NADP IS COUPLED

```

```

CC TO INSPIRATION AND ATP HYDROLYSIS AND FUNCTIONS AS A PROTON PUMP
CC ACROSS THE MEMBRANE.
CC -1- CATALYTIC ACTIVITY: NADPH + NAD(+) = NADP(+) + NADH.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: TO THE N-TERMINAL PART OF THE EUKARYOTIC MITOCHONDRIAL
CC ENZYME AND TO ALANINE DEHYDROGENASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04195; CAB37089.1; -
CC EMBL; X66086; CAA46884.1; -
CC EMBL; AE000255; AAC74675.1; -
CC EMBL; D90803; BAA15342.1; -
CC EMBL; D90804; BAA15352.1; -
CC EMBL; D90802; BAA15337.1; -
CC PIR; A25012; DECCA.
CC PIR; S24380; S24380.
CC PIR; S23833; S23833.
CC HSSP; Q60164; 1H2Z.
CC Ecogene; EG10744; pntA.
CC Interpro; IPR004002; Aladh_PNT.
CC Interpro; IPR004571; pntA.
CC Pfam; PF01262; Aladh_PNT; 1.
CC TIGRfams; TIGR00561; pntA; 1.
CC PROSITE; PS00836; ALADH_PNT_1; 1.
CC PROSITE; PS00837; ALADH_PNT_2; 1.
CC Oxidoreductase; NAD; NADP; Transmembrane; Inner membrane;
CC Complete proteome.
CC KW NP-BIND 167 NAD (By SIMILARITY).
CC FT TRANSMEM 402 422 POTENTIAL.
CC FT TRANSMEM 423 443 POTENTIAL.
CC FT TRANSMEM 453 473 POTENTIAL.
CC FT TRANSMEM 477 497 POTENTIAL.
CC SQ SEQUENCE 510 AA; 54623 MW; 801742097BEA6943 CRC64;

Query Match 40.7%; Score 48; DB 1; Length 510;
Best Local Similarity 52.9%; Pred. No. 6.5;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0.

QY 6 FRGIVHGRSIRGFTG 22
   :||| |||
Db 139 YRAIVEAHHEGRFTG 155

RESULT 3
CYB_MAIZE STANDARD: PRT: 388 AA.
CID CYB_MAIZE
PC P04165;
DT 01-AUG-1987 (Rel. 04 Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Zea mays (Maize).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Dawson A.J., Jones V.P., Leaver C.J.;
RT "The apocytochrome b gene in maize mitochondria does not contain
RT introns and is preceded by a potential ribosome binding site.";
RL EMBD J. 3:2107-2113(1984).
CL -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

```

```

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X00789; CAA25367.1; -.
CC PIR: A00156; CBZM.
CC MaizelDB: 69227; -.
DR InterPro: IPR000179; Cyt_b_b6.
DR Pfam: PF00032; cytochrome_b_c1.
DR Pfam: PF00033; cytochrome_b_n; 1.
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE: PS00193; CYTOCHROME_B_OO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 88 88 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 102 102 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 189 189 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 203 203 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 388 AA: 43567 MW: 56CF734C2B62B3F6 CRC64;

Query Match 38.6%; Score 45.5; DB 1; Length 388;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

OY 2 IH-HIFRGIVHAGRSIGR 18
DB 99 VHLHIFRGLYHASYSPPR 116

RESULT 4
CYB_SOLTM
ID CYB_SOLTM STANDARD; PRT: 392 AA.
AC P29757;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Solanum tuberosum (Potato).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Bintje; PubMed=1840690;
RX MEDLINE=91355947;
RA Zaulungo S., Litvak S., Jordana X.;
RT "Isolation and nucleotide sequence of the potato mitochondrial gene
RT for apocytochrome b."
RL Plant Mol. Biol. 17:527-530(1991).
RN [2]
RN SEQUENCE OF 1-18.
RP TISSUE=Tuber;
RC MEDLINE=93131029; PubMed=8420797;
RA Braun H.-P., Schmitz U.K.;
RT "Purification and sequencing of cytochrome b from potato reveals
RT methionine cleavage of a mitochondrially encoded protein."
RL FEBS Lett. 316:128-132(1993).

```

```

CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- PFM: FIRST MITOCHONDRIAL ENCODED PROTEIN TO BE SHOWN TO HAVE ITS
CC N-TERMINAL METHIONINE CLEAVED OFF.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X58437; CAA41343.1; -.
CC PIR: S17427; CBPM.
DR InterPro: IPR000179; Cyt_b_b6.
DR Pfam: PF00032; cytochrome_b_c1.
DR Pfam: PF00033; cytochrome_b_n; 1.
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE: PS00193; CYTOCHROME_B_OO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT INIT_MET 0 0
FT METAL 87 87 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 101 101 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 188 188 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 202 202 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 392 AA: 43913 MW: CA147084333F73AC CRC64;

Query Match 38.6%; Score 45.5; DB 1; Length 392;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

OY 2 IH-HIFRGIVHAGRSIGR 18
DB 98 VHLHIFRGLYHASYSPPR 115

RESULT 5
CYB_VICFA
ID CYB_VICFA STANDARD; PRT: 392 AA.
AC P05718;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
GN COB OR CYTB.
OS Vicia faba (Broad bean).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX NCBI_TaxID=3906;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88303319; PubMed=3405753;
RA Wahleithner J.A., Wolstenholme D.R.;
RT "Ribosomal protein S14 genes in broad bean mitochondrial DNA."
RL Nucleic Acids Res. 16:6897-6913(1988).
RN [2]
RN FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,

```

CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X07237; CAA30226.1; -
 CC PIR: S01221; CBVF.
 CC InterPro: IPR000179; Cyt_b_b6.
 CC Pfam: PF00032; cytochrome_b_c_1.
 CC PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
 CC PROSITE: PS00193; CYTOCHROME_B_OO; 1.
 CC Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
 KW Heme.
 FT METAL 88 88 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 102 102 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 189 189 IRON 2 (HEME B562 AXIAL LIGAND).
 FT METAL 203 203 IRON 1 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 392 AA; 44013 MW; 023D3396B2FA2BD3 CRC64;
 Query Match 38.6%; Score 45.5; DB 1; Length 392;
 Best Local Similarity 55.6%; Pred. No. 12;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 QY 2 IH-HIFRGIVHAGRSIGR 18
 Db 99 VHLHIFRGIVHAGRSIGR 116

CC 329 AND 363 ARE MODIFIED BY RNA EDITING.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X07126; CAA30135.1; -
 CC PIR: S20141; CROBE.
 CC InterPro: IPR000179; Cyt_b_b6.
 CC Pfam: PF00032; cytochrome_b_c_1.
 CC PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
 CC PROSITE: PS00193; CYTOCHROME_B_OO; 1.
 CC Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
 KW Heme; RNA editing.
 FT METAL 89 89 IRON 1 (HEME B562 AXIAL LIGAND).
 FT METAL 103 103 IRON 2 (HEME B566 AXIAL LIGAND).
 FT METAL 191 191 IRON 2 (HEME B562 AXIAL LIGAND).
 FT METAL 204 204 IRON 1 (HEME B566 AXIAL LIGAND).
 SQ SEQUENCE 394 AA; 44152 MW; D8CAE874866C20FA CRC64;
 Query Match 38.6%; Score 45.5; DB 1; Length 394;
 Best Local Similarity 55.6%; Pred. No. 12;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 QY 2 IH-HIFRGIVHAGRSIGR 18
 Db 100 VHLHIFRGIVHAGRSIGR 117

RESULT 7
 CYB_ORYSA
 ID CYB_ORYSA STANDARD; PRT; 397 AA.
 AC P14833;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome B.
 GN COB OR CYTB.
 OS Oryza sativa (Rice).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Indica-IR36;
 RX MEDLINE=90221830; PubMed=2326174;
 RA Kalelkar E.K., Andre C.P., Doshi B., Walbot V.;
 RT "Sequence of the rice mitochondrial gene for apocytochrome b.";
 RL Nucleic Acids Res. 18:372-372(1990).
 CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
 CC COUPLED TO ATP SYNTHESIS.
 CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
 CC BOUND TO THE PROTEIN.
 CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>


```

CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: X17064; CAA34910.1; -
CC PIR: J00164; CBR.
CC InterPro: IPR000179; Cyt_b_b6.
CC Pfam: PF00032; cytochrome_b_c_1.
CC DR PFam: PF00033; cytochrome_b_n_1.
CC DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
CC DR PROSITE: PS00193; CYTOCHROME_B_OO; 1.
CC KM Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
CC Heme.
CC METAL 88 88 IRON 1 (HEME B562 AXIAL LIGAND).
CC METAL 102 102 IRON 2 (HEME B566 AXIAL LIGAND).
CC METAL 189 189 IRON 2 (HEME B562 AXIAL LIGAND).
CC METAL 203 203 IRON 1 (HEME B566 AXIAL LIGAND).
CC FT SEQUENCE 397 AA; 44540 MW; D9122B9FC20D6DE CRC64;
CC
CC Query Match 38.6%; Score 45.5; DB 1; Length 397;
CC Best Local Similarity 53.6%; Pred. No. 12;
CC Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
CC
CC QY 2 IH-HIFRGIVHAGRSIGR 18
CC Db 99 VHLHIFRGIVHAGRSYSSPR 116
CC
CC RESULT 8
CC PID4_MOUSE STANDARD; PRT; 666 AA.
CC AC 092183;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Protein-arginine deiminase type IV (EC 3.5.3.15) (Peptidylarginine
CC deiminase IV).
CC GN PAD14 OR PID4.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC RX NCBI_TaxID=10090;
CC RX [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Epidermis;
CC RX MEDLINE=99192810; PubMed=10092850;
CC RX Rusd A.A., Ikejiri Y., Ono H., Yonekawa T., Shiraiwa M., Kawada A.,
CC RA Takahara H.;
CC RT "Molecular cloning of cDNAs of mouse peptidylarginine deiminase type
CC RT I, type III and type IV, and the expression pattern of type I in
CC RT mouse".
CC RL Eur. J. Biochem. 259:660-669(1999).
CC CC -I- FUNCTION: CATALYZES THE DEIMINATION OF ARGININE RESIDUES OF
CC CC PROTEINS.
CC CC -I- CATALYTIC ACTIVITY: Protein L-arginine + H(2)O = protein L-
CC CC citrulline + NH(3).
CC CC -I- COFACTOR: REQUIRES CALCIUM IONS.
CC CC -I- SIMILARITY: SOME SIMILARITY WITH ARGINASE AND TO ARGININE
CC CC DEIMINASE.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@lsb-sib.ch).
CC CC -----
CC CC EMBL: AB013850; BAA34246.1; -
CC CC MGD: MGI:1338898; Pq14.
CC CC InterPro: IPR004303; Prot_arg_deim.
CC DR Pfam: PF03068; PAD; 1.
CC KM Hydrolase; Calcium-binding; Multigene family.
CC FT CA_BIND 505 516 EF-HAND (POTENTIAL).
CC SEQUENCE 666 AA; 74476 MW; 70FAE4E7E232D3A4 CRC64;

```

```

CC Query Match 38.1%; Score 45; DB 1; Length 666;
CC Best Local Similarity 33.3%; Pred. No. 24;
CC Matches 6; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 1 FTHHIFRGIVHAGRSIGR 18
CC Db 637 YTHHIFRGIVHAGRSIGR 654
CC
CC RESULT 9
CC SYRM_RHIEF STANDARD; PRT; 336 AA.
CC AC 008812;
CC DT 01-FEB-1995 (Rel. 31, Created)
CC DT 01-FEB-1995 (Rel. 31, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE SYRM protein (Symbiotic regulator).
CC GN SYRM.
CC OS Rhizobium etli.
CC OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC OC Rhizobiaceae; Rhizobium.
CC RX NCBI_TaxID=29449;
CC RX [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=CNPAF512;
CC RX MEDLINE=93376517; PubMed=8367305;
CC RX Michiels J., de Wilde P., Vanderleyden J.;
CC RT "Sequence of the Rhizobium leguminosarum biovar phaseoli syrm gene.";
CC RL Nucleic Acids Res. 21:3893-3893(1993).
CC CC -I- FUNCTION: SYRM ACTS IN TRANS TO STIMULATE NOD GENE EXPRESSION VIA
CC CC NODD3 AND EXO GENE EXPRESSION VIA SYRM.
CC CC -I- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC CC REGULATORS.
CC CC -I- CAUTION: STRAIN CNPAF512 WAS ORIGINALLY THOUGHT TO ORIGINATE FROM
CC CC R. LEGUMINOSARUM BV. PHASEOLI.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@lsb-sib.ch).
CC CC -----
CC CC EMBL: Z23013; CAA80556.1; -
CC CC InterPro: IPR000847; HTH_LysR.
CC DR InterPro: IPR005119; LysR_subst.
CC DR Pfam: PF00126; HTH_1; 1.
CC DR Pfam: PF03466; LysR_substrate; 1.
CC DR PROSITE: PS00044; HTH_LysR_FAMILY; 1.
CC KW Modulation; Transcription regulation; DNA-binding; Activator.
CC FT DNA_BIND 58 77 H-T-H MOTIF (BY SIMILARITY).
CC SEQUENCE 336 AA; 37800 MW; FACE9463A6A872BD CRC64;
CC
CC Query Match 37.3%; Score 44; DB 1; Length 336;
CC Best Local Similarity 53.8%; Pred. No. 17;
CC Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 6 FRGIVHAGRSIGR 18
CC Db 55 YRNVTHAGOHIGR 67
CC
CC RESULT 10
CC SYM_HALNT STANDARD; PRT; 380 AA.
CC AC 09HN66;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
CC (TrpKs).

```

GN TRPS OR TRPS2 OR VNG2232G.
 OS Halobacterium sp. (strain NRC-1).
 CC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
 CC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20504483; PubMed=11016950;
 RA NG W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shroog J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leitner B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Fohlschoder M., Spatich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Elmhart H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
 RT "Genome sequence of Halobacterium species NRC-1".
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC dihydrophosphate + L-tryptophanyl-tRNA(Trp).
 CC SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE005109; AAC20355.1;
 DR InterPro: IPR002306; Trp-trna-synt_1b.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PRO1039; TRNASYNTTRP.
 DR TIGRFAMS: TIGR00233; trps; 1.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_I; 1.
 KM Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding;
 KW Complete proteome.
 FT SITE 74 "HIGH" REGION.
 FT SITE 249 "KMSKS" REGION.
 FT SITE 253 "KMSKS" REGION.
 SO SEQUENCE 380 AA; 41936 MW; 0F7B2B955386404F CRC64;
 Query Match 37.3%; Score 44; DB 1; Length 380;
 Best Local Similarity 53.3%; Pred. No. 20;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 OY 7 RGIIVHGRSIGRFLT 21
 Db 48 RGIIVHGRSIGRFLT 62
 RESULT 11
 PSAB_HETTR STANDARD; PRT; 776 AA.
 ID PSAB_HETTR
 AC Q9XOV2;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Photosystem I P700 chlorophyll A apoprotein A2 (Psab) (PSI-B).
 GN PSAB.
 OS Heterocapsa triquetra (Dinoflagellate).
 OC Chloroplast.
 OC Eukaryota; Alveolata; Dinophyceae; Peridinales; Heterocapsaceae;
 OC Heterocapsa.
 OX NCBI_TaxID=66468;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CCMP 449;
 RX MEDLINE=99334925; PubMed=10408440;
 RX Zhang Z., Green B.R., Cavalier-Smith T.;

RT "Single gene circles in dinoflagellate chloroplast genomes."
 RL Nature 400:155-159(1999).
 CC -1- FUNCTION: Psab and psab bind P700, the primary electron donor of
 CC photosystem I (PSI), as well as the electron acceptors A0, A1, and
 CC FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin
 CC oxidoreductase.
 CC COFACTOR: P700 is a chlorophyll A dimer. A0 is chlorophyll A, A1
 CC is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.
 CC -1- SUBUNIT: A psab/psab heterodimer binds the P700 chlorophyll special
 CC pair and subsequent electron acceptors. The PSI reaction center of
 CC higher plants and algae is composed of one at least 11 subunits.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane.
 CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF130032; AAD4699.1;
 DR HSSP: P25897; JUBO.
 DR InterPro: IPR001280; PSI_Psaa/B.
 DR Pfam: PF00223; psaa-psab; 1.
 DR PRINTS: PRO0257; PHOTOSPSAB.
 DR PROSITE: PS00419; PHOTOSYSTEM_I_PSAAB; 1.
 KM Photosynthesis; Photosystem I; Transport; Electron transport;
 KM Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;
 KW Iron-sulfur; 4Fe-4S; Chlorophyll.
 FT TRANSMEM 57 80 I (POTENTIAL).
 FT TRANSMEM 150 172 II (POTENTIAL).
 FT TRANSMEM 233 257 III (POTENTIAL).
 FT TRANSMEM 334 352 IV (POTENTIAL).
 FT TRANSMEM 382 405 V (POTENTIAL).
 FT TRANSMEM 421 447 VI (POTENTIAL).
 FT TRANSMEM 472 494 VII (POTENTIAL).
 FT TRANSMEM 561 579 VIII (POTENTIAL).
 FT TRANSMEM 619 640 IX (POTENTIAL).
 FT TRANSMEM 687 709 X (POTENTIAL).
 FT TRANSMEM 753 773 XI (POTENTIAL).
 FT METAL 603 603 IRON-SULFUR FX (4Fe-4S) (BY SIMILARITY).
 FT METAL 612 612 IRON-SULFUR FX (4Fe-4S) (BY SIMILARITY).
 FT METAL 698 698 P700 SPECIAL PAIR CHLOROPHYLL AXIAL
 FT BINDING 698 698 LIGAND (BY SIMILARITY).
 FT BINDING 706 706 A0 CHLOROPHYLL (BY SIMILARITY).
 FT BINDING 714 714 A1 PHYLLOQUINONE (BY SIMILARITY).
 FT BINDING 715 715 A1 PHYLLOQUINONE (BY SIMILARITY).
 FT BINDING 746 746 A1 PHYLLOQUINONE (BY SIMILARITY).
 SO SEQUENCE 776 AA; 86143 MW; F8640CA5F58C8913 CRC64;
 Query Match 37.3%; Score 44; DB 1; Length 776;
 Best Local Similarity 50.0%; Pred. No. 40;
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 OY 3 HIRFGIVHGRSIGRFLTG 22
 Db 302 HIRFGIVHGRSIGRFLTG 321
 RESULT 12
 ARP2_CAEEL STANDARD; PRT; 395 AA.
 ID ARP2_CAEEL
 AC P53489;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Actin-like protein 2 (Actin-like protein C).
 GN ACTC OR K07C5.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Chordata; Rhabditida; Rhabditidae;

```

OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA McMurtry A.;
RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PART OF A COMPLEX IMPLICATED IN THE CONTROL OF ACTIN
CC POLYMERIZATION IN CELLS (BY SIMILARITY).
CC -1- SUBUNIT: BELONGS TO A COMPLEX COMPOSED OF ARP2, ARP3, P41-ARC,
CC P34-ARC, P21-ARC, P20-ARC AND P16-ARC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ACTIN FAMILY. ARP2 SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z71181; CA94894.1; -
DR WormPep; K07C5.1; CE06111.
DR InterPro; IPR004001; Actin.
DR InterPro; IPR004000; Actin_Like.
DR Pfam; PF00022; actin; 1.
DR SMART; SM00268; ACTIN; 1.
DR PROSITE; PS00432; ACTINS_2; 1.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
DR Structural protein; Cytoskeleton.
KW SDOJNCE 395 AA; 44840 MW; D13B735000420692 CRC64;
SQ
Query Match 36.4%; Score 43; DB 1; Length 395;
Best Local Similarity 42.1%; Pred. No. 29;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
OY 2 IHHIFRGIVHAGRSIGRFL 20
DB 175 LHHLTRLDIAGRDITKYL 193
::: | : | | | : | |
RESULT 13
ID NSMA_CAEEL STANDARD; PRT; 445 AA.
AC 045870;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative neutral sphingomyelinase (EC 3.1.4.12).
GN T27F6.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Dobson R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Sphingomyelin + H(2)O = N-acylsphingosine +
CC choline phosphate.
CC -1- SIMILARITY: BELONGS TO THE NEUTRAL SPHINGOMYELINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z82060; CAB04885.1; -
DR WormPep; T27F6.6; CE16515.

```

```

DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
KW Hypothetical protein; Hydrolase; Transmembrane.
FT TRANSMEM 377 399
FT TRANSMEM 377 399
FT METAL 403 425
FT METAL 83 83
FT SITE 215 215
FT ACT_SITE 318 318
SQ SEQUENCE 445 AA; 51054 MW; 969FEC3A037FE72 CRC64;
OY 1 FIHHIFRGIVHAGRSIG 17
DB 137 FAHHIHGDMFGCKVVG 153
::: | | | | : | |
RESULT 14
ID GVD2_HALNI STANDARD; PRT; 492 AA.
AC 09HHI2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GVPD Protein 2.
GN GVPD2 OR VNG6240G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxId=64091;
RM [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950.
RA Ng W.V., Kennedy S.P., Mahairs G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
RA Leitauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Isenbarger T.A., Peck R.F., Pohlschoder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- FUNCTION: MAY PLAY A ROLE IN AN ENERGY REQUIRING PROCESS SUCH AS
CC ASSEMBLY OF GAS VESICLES IN ADDITION TO A POSSIBLE STRUCTURAL OR
CC REGULATORY FUNCTION.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE005156; AAG20894.1; -
DR Gas vesicle; ATP-binding; Plasmid; Complete proteome.
KW NP_BIND 39 46
FT SEQUENCE 492 AA; 53928 MW; 655965FD1C3631B CRC64;
OY 8 GIYHAGSICGRFLTG 22
DB 15 GIAHFPRIRRFETG 29
::: | | | | : | |
Query Match 36.4%; Score 43; DB 1; Length 492;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

RESULT 15
GVD2_HALN2      STANDARD:      PRT:      492 AA.
ID      GVD2_HALN2
AC      P33958;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      GvpD protein 2.
GN      GVPD OR GVPD.
OS      Halobacterium sp. (strain NRC-817).
OC      Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC      Halobacteriaceae; Halobacterium.
OX      NCBI_TaxID=148370;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93021102; PubMed=1404376;
RA      Engleert C., Krueger K., Offner S., Pfeifer F.;
RT      "Three different but related gene clusters encoding gas vesicles in
RT      halophilic archaea.";
RL      J. Mol. Biol. 227:586-592(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96312339; PubMed=8763925;
RA      Krueger K., Pfeifer F.;
RT      "Transcript analysis of the c-vac region and differential synthesis of
RT      the two regulatory gas vesicle proteins GvpD and Gvpe in Halobacterium
RT      salinarium PH4.";
RL      J. Bacteriol. 178:4012-4019(1996).
CC      -1- FUNCTION: MAY PLAY A ROLE IN AN ENERGY REQUIRING PROCESS SUCH AS
CC      ASSEMBLY OF GAS VESICLES IN ADDITION TO A POSSIBLE STRUCTURAL OR
CC      REGULATORY FUNCTION.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X64730; CAA45992.1; -.
DR      EMBL; X94688; CAA64343.1; -.
DR      PIR; S28139; S28139.
KM      Gas vesicle; ATP-binding.
FT      NP_BIND 39
SQ      SEQUENCE 492 AA; 54100 MW; 3D30F6E317ABA34E CRC64;

Query Match      36.4%; Score 43; DB 1; Length 492;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY      8 GIVHAGRSIGRLTG 22
      || | | | | | |
DB      15 GIAHFPREIRRFETG 29

```

Search completed: May 17, 2003, 14:25:41
 Job time : 27 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2003, 14:21:19 ; Search time 83 Seconds
(without alignments)
54.615 Million cell updates/sec

Title: US-09-929-788-1

Perfect score: 118
Sequence: 1 FHHFRGIVHAGRSIGRLTG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_proteint:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	73.7	79	13	08UUG0
2	84	71.2	79	13	08UUG2
3	53	44.9	523	2	Q9ALA2
4	53	44.9	524	16	Q8UCJ1
5	49	41.5	205	16	Q985H7
6	49	41.5	763	10	Q9FR59
7	48	40.7	83	16	08UB08
8	48	40.7	508	16	08ZEF9
9	48	40.7	509	16	08ZEP0
10	48	40.7	509	16	08ZEX5
11	48	40.7	883	16	Q9CHA3
12	48	40.7	887	12	Q9YW93
13	48	40.7	894	12	Q9YW92
14	47.5	40.3	357	16	08UDH4
15	47	39.8	273	2	066090
16	47	39.8	512	16	Q9KM26

17	47	39.8	513	16	Q9JZNO	Q9JZNO neisseria m
18	47	39.8	513	16	Q9JUR3	Q9JUR3 neisseria m
19	47	39.8	1762	2	Q52546	Q52546 amycolatops
20	47	39.8	1763	2	Q52790	Q52790 amycolatops
21	46	39.0	360	8	Q47575	Q47575 onchocerca
22	46	39.0	480	5	Q9VU45	Q9VU45 drosophila
23	46	39.0	517	5	Q8T063	Q8T063 drosophila
24	46	39.0	530	16	P73496	P73496 synechocyst
25	46	39.0	537	10	Q9FW78	Q9FW78 oryza sativ
26	46	39.0	579	2	Q67993	Q67993 pseudomonas
27	46	39.0	687	10	Q9LMT8	Q9LMT8 arabidopsis
28	46	39.0	767	17	Q9HID7	Q9HID7 thermoplasma
29	46	39.0	880	16	Q8YXN7	Q8YXN7 anabaena sp
30	45.5	38.6	212	8	Q9T5A5	Q9T5A5 solanum tub
31	45.5	38.6	251	16	Q8XW76	Q8XW76 ralsionia s
32	45.5	38.6	291	8	Q35221	Q35221 oenothera b
33	45.5	38.6	392	8	Q9ZET8	Q9ZET8 pisum sativ
34	45.5	38.6	333	8	Q9S871	Q9S871 nicotiana t
35	45.5	38.6	333	8	Q9MF98	Q9MF98 beta vulgar
36	45.5	38.6	397	8	Q35293	Q35293 oryza sativ
37	45.5	38.6	424	8	Q9B3G0	Q9B3G0 triticum ae
38	45.5	38.6	2731	11	Q9WTS4	Q9WTS4 mus musculu
39	45	38.1	140	5	Q25277	Q25277 leishmania
40	45	38.1	141	2	Q31022	Q31022 vibrio chol
41	45	38.1	141	16	Q9KMH2	Q9KMH2 vibrio chol
42	45	38.1	182	13	Q90YG5	Q90YG5 oncorhynch
43	45	38.1	260	17	Q8TUN6	Q8TUN6 methanosarc
44	45	38.1	321	17	Q973G7	Q973G7 sulfolobus
45	45	38.1	388	17	Q96XU4	Q96XU4 sulfolobus

ALIGNMENTS

RESULT 1

Q8UUG0 ID Q8UUG0 PRELIMINARY; PRT; 79 AA.
AC Q8UUG0;
DT 01-MAR-2002 (TREMREL. 20, Created)
DT 01-MAR-2002 (TREMREL. 20, Last sequence update)
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)
DE Moronecidin.
OS Morone saxatilis (Striped bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percioidei;
OC Moronidae; Morone.
OX NCBI_TaxID=34816;
RN [1]
RP SEQUENCE FROM N.A.
RA Lauth X., Shike H., Burns J.C., Westerman M.E., Ostland V.E.,
RA Carlberg J.M., Van Olst J.C., Nizet V., Taylor S.W., Shimizu C.,
RA Bulet P.;
RT "Discovery and characterization of two isoforms of moronecidin, a
RT novel antimicrobial peptide from hybrid striped bass."
RL J. Biol. Chem. 0:0-0(2002).
DR EMBL: AF394244; AAL57319.1; -;
DR EMBL: AF385583; AAL49496.1; -;
SQ SEQUENCE 79 AA: 9222 MW; C9DF13E1DAFA7EFD CRC64;

Query Match 73.7%; Score 87; DB 13; Length 79;
Best Local Similarity 68.2%; Pred. No. 4; Ee-06;
Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FHHFRGIVHAGRSIGRLTG 22
DB 23 FHHFRGIVHAGRSIGRLTG 44

RESULT 2
Q8UUG2 ID Q8UUG2 PRELIMINARY; PRT; 79 AA.
AC Q8UUG2;

DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE Moronecidin prepropeptide precursor.
 OS Morone chrysopt (White Bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 NCBI_TaxID=46259;
 RN
 RP SEQUENCE FROM N.A.
 RA Lauch X.C., Shike H., Burns J.C., Westernman M.E., Ostland V.E.,
 RA Carlsberg J.M., Van Olst J.C., Nizet V., Taylor S.W., Shlimizu C.,
 RA Bulet P.;
 RT "Discovery, and characterization of two isoforms of moronecidin, a
 RT novel antimicrobial peptide from hybrid striped bass.";
 RL J. Biol. Chem. 0:0-0(2002).
 DR EMBL: AF333621; AAL40409.1; -;
 DR EMBL: AF394243; AAL57318.1; -;
 KW Signal.
 FT SIGNAL. 1 22 POTENTIAL.
 FT CHAIN 23 44 MORONECIDIN.
 SQ SEQUENCE 79 AA; 9106 MW; B6F32481C4200EEB CRC64;

Query Match 71.2%; Score 84; DB 13; Length 79;
 Best Local Similarity 63.6%; Pred. No. 1.3e-05;
 Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FHHHFRGIVHAGRSIGRFLTG 22
 1 ||||| : : : : :
 Db 23 FFHHFRGIVHAGRSIGRFLTG 44

RESULT 3
 Q9ALA2 PRELIMINARY; PRT; 523 AA.
 AC Q9ALA2;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Pyridine nucleotide transhydrogenase alpha subunit.
 GN PNTA.
 OS Rhodobacter sphaeroides (Rhodospirillum rubrum).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Rhodobacter.
 NCBI_TaxID=1063;
 RN
 RP SEQUENCE FROM N.A.
 RA Hickman J.W., Barber R.D., Skaar E., Donohue T.J.;
 RT "A role for pyridine nucleotide transhydrogenase in formaldehyde
 RT metabolism in Rhodobacter sphaeroides.";
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY026033; AAK00588.1; -;
 DR HSSP: Q60164; 1H2Z.
 DR InterPro: IPR004571; PNTA.
 DR Pfam: PF01262; Aladh_PNT; 1.
 DR TIGRfams: TIGR00561; PNTA; 1.
 GN
 SQ SEQUENCE 523 AA; 55618 MW; 45DFCF75683EF1C9 CRC64;

Query Match 44.9%; Score 53; DB 2; Length 523;
 Best Local Similarity 47.1%; Pred. No. 6.3;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 6 FRGIVHAGRSIGRFLTG 22
 : : : : :
 Db 141 YRAVIEAGNNGFRFTG 157

RESULT 4
 ID 08UCJ1 PRELIMINARY; PRT; 524 AA.
 AC 08UCJ1;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE NAD(P)+ transhydrogenase.
 GN PNTA OR AFU2493 OR AGR_C_4529.
 OS Agrobacterium tumefaciens (Strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 NCBI_TaxID=176299;
 RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE-21608550; PubMed-11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gilliet W., Grant C.,
 RA Kutayian T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Neeter E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE-21608551; PubMed-11743194;
 RA Gordon B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Ounullo B., Goldman B.S., Cao Y., Askew M., Halling C., Millin L.,
 RA Houmlet K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
 RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009196; AAL43480.1; -;
 DR EMBL: AE008162; AAK8225.1; -;
 KW Complete proteome.
 SQ SEQUENCE 524 AA; 55862 MW; 581947CD3BD6B87 CRC64;

Query Match 44.9%; Score 53; DB 16; Length 524;
 Best Local Similarity 47.1%; Pred. No. 6.3;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 6 FRGIVHAGRSIGRFLTG 22
 : : : : :
 Db 141 YRAVIEAGNNGFRFTG 157

RESULT 5
 ID 0985H7 PRELIMINARY; PRT; 205 AA.
 AC 0985H7;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE Hypothetical protein ml17669.
 GN ML17669.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Pyllobacteriaceae; Mesorhizobium.
 NCBI_TaxID=381;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idehara K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003012; BAB54085.1; -
DR InterPro: IPR001387; HTH_3.
DR Pfam: PF01381; HTH_3; 1.
DR SMART: SM00530; HTH_XRE; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 205 AA; 22345 MW; 04F25BE0973ACC38 CRC64;

Query Match
Best Local Similarity 41.5%; Score 49; DB 16; Length 205;
Matches 12; Conservative 1; Mismatches 8; Indels 6; Gaps 1;

QY 2 IHIFRG-----IVHAGRSIGRFLTG 22
DB 131 IHIEAGSGAGPIQHAGEVGYLLEG 157

RESULT 6
Q9FR59 PRELIMINARY; PRT; 763 AA.
ID Q9FR59;
AC Q9FR59;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
DE Homeobox 1.
GN HBI.
OS Picea abies (Norway spruce) (Picea excelsa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
OX NCBI_TaxID=3329;
RN [1]
RP SEQUENCE FROM N.A.
RA Ingouff M., Farbos I., von Arnold S.;
RT "Pahb1, a HD-GU2 family homeobox gene, is expressed in the outer cell
RT layer during somatic embryogenesis in Norway spruce.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AE172931; AAG43405.1; -
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR000047; HTH_repressor.
DR InterPro: IPR002913; START.
DR Pfam: PF00046; homeobox; 1.
DR Pfam: PF01852; START; 1.
DR PRINTS: PR00031; HTHREPRESSR.
DR PRODOM: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR SMART: SM00234; START; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS00071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 763 AA; 83795 MW; 0EF8D5851427BA2 CRC64;

Query Match
Best Local Similarity 41.5%; Score 49; DB 10; Length 763;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 IHIFRGIVHAGRSIG 17
DB 472 VHHIYRLVNSGMAFG 487

RESULT 7

```

```

Q8UB08
ID Q8UB08 PRELIMINARY; PRT; 83 AA.
AC Q8UB08;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Hypothetical protein Atu3209.
GN Atu3209.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiales group;
OC Rhizobiales; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Katp P.D., Hovee D., Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
DR EMBL: AE009250; AL44025.1; -
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 83 AA; 9062 MW; 5392C9518F4A0CA CRC64;

Query Match
Best Local Similarity 40.7%; Score 48; DB 16; Length 83;
Matches 11; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

QY 2 IHIFRGIVHAGRSIGRFLTG 22
DB 10 VHH-----GAGRRIGRFLTG 24

RESULT 8
Q8ZE79 PRELIMINARY; PRT; 508 AA.
ID Q8ZE79;
AC Q8ZE79;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2).
GN PNTA OR YP02303.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tariga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan A.V.,
RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL: AJ414151; CAC91108.1; -
DR InterPro: IPR004002; AlAdh_PNT.
DR InterPro: IPR004571; PNTA.
DR Pfam: PF01262; AlAdh_PNT; 1.
DR TIGRFAMS: TIGR00561; pnta; 1.
DR PROSITE: PS00836; ALADH_PNT_1; 1.

```

DR PROSITE: PS00837; ALADH_PNT_2; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 508 AA; 54270 MW; 8198AC8BD82411F CRC64;

Query Match 40.7%; Score 48; DB 16; Length 508;
 Best Local Similarity 52.9%; Pred. No. 36;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIYHAGRSIGRFLTG 22
 :| | | | | | | | | |
 Db 139 YRAIVEAHEHGRFFETG 155

RESULT 9
 08ZPL0 PRELIMINARY; PRT; 509 AA.

AC 08ZPL0;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DE Pyridine nucleotide transhydrogenase (proton pump), alpha subunit (EC 1.6.1.2).
 GN PNTA OR STM1479.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.

RC MEDLINE=21534948; PubMed=11677609;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewel N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 DR EMBL; AE008764; AAL20398.1; -;
 DR InterPro: IPR004002; ALADH_PNT.
 DR InterPro: IPR004571; PntA.
 DR Pfam: PF01262; ALADH_PNT_1;
 DR TIGRFAMS: TIGR00561; pntA; 1.
 DR PROSITE: PS00836; ALADH_PNT_1; 1.
 DR PROSITE: PS00837; ALADH_PNT_2; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 509 AA; 54266 MW; ECC3JAB4537C0D2E CRC64;

Query Match 40.7%; Score 48; DB 16; Length 509;
 Best Local Similarity 52.9%; Pred. No. 37;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIYHAGRSIGRFLTG 22
 :| | | | | | | | | |
 Db 139 YRAIVEAHEHGRFFETG 155

RESULT 10
 08ZKX5 PRELIMINARY; PRT; 509 AA.

AC 08ZKX5;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DE Pyridine nucleotide transhydrogenase subunit-alpha.
 GN STY1589.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moulie S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 DR EMBL; AL627270; CAD01835.1; -;
 DR InterPro: IPR004002; ALADH_PNT.
 DR InterPro: IPR004571; PntA.
 DR Pfam: PF01262; ALADH_PNT_1;
 DR TIGRFAMS: TIGR00561; pntA; 1.
 DR PROSITE: PS00836; ALADH_PNT_1; 1.
 DR PROSITE: PS00837; ALADH_PNT_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 509 AA; 54236 MW; OCC39004537C0D34 CRC64;

Query Match 40.7%; Score 48; DB 16; Length 509;
 Best Local Similarity 52.9%; Pred. No. 37;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 FRGIYHAGRSIGRFLTG 22
 :| | | | | | | | | |
 Db 139 YRAIVEAHEHGRFFETG 155

RESULT 11
 09CHA3 PRELIMINARY; PRT; 883 AA.

ID 09CHA3;
 AC 09CHA3;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein yieh.
 GN YIEH OR L10835.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=ILI1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
 RA Weisenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403.";
 RL Genome Res. 11:731-753(2001).
 DR EMBL; AE006317; AAK04933.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 883 AA; 99524 MW; 42ADC278C1FDC191 CRC64;

Query Match 40.7%; Score 48; DB 16; Length 883;
 Best Local Similarity 45.0%; Pred. No. 65;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 IHIIHFIYHAGRSIGRFLT 21
 :| | | | | | | | | |
 Db 51 IHSIVNIIHSGSGSGLT 70

RESULT 12
 09YW93 PRELIMINARY; PRT; 887 AA.

ID 09YW93;
 AC 09YW93;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Glycoprotein B.
 GN UL27.
 OS Macropodid herpesvirus type 1 (parma wallaby herpesvirus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC unclassified Herpesviridae.
 OX NCBI_TaxID=83441;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9171712; PubMed=10073704;
 RA Mahony T.J., Smith G.A., Thomson D.M.;
 RT "Macropodid herpesviruses 1 and 2 occupy unexpected molecular
 phylogenetic positions within the Alphaherpesvirinae.";
 RL J. Gen. Virol. 80:433-436(1999).
 DR EMBL; AF061754; AAD11960.1;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein B.1.
 DR ProDom; PD000693; Glycoprot_B.1.
 SQ SEQUENCE 887 AA; 99739 MW; 0EE4FB4E098F128 CRC64;

Query Match 40.7%; Score 48; DB 12; Length 887;
 Best Local Similarity 38.1%; Pred. No. 65;
 Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 2 IHHRGIVHAGRSIGRPLTG 22
 Db 718 IYSPFDGIGIGRAVGRVVTG 738

RESULT 13

OY9W92 PRELIMINARY; PRT; 894 AA.
 AC O9YW92;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Glycoprotein B.
 GN UL27.
 OS Macropodid herpesvirus type 2 (dorcopsis wallaby herpesvirus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC unclassified Herpesviridae.
 OX NCBI_TaxID=83440;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99171712; PubMed=10073704;
 RA Mahony T.J., Smith G.A., Thomson D.M.;
 RT "Macropodid herpesviruses 1 and 2 occupy unexpected molecular
 phylogenetic positions within the Alphaherpesvirinae.";
 RL J. Gen. Virol. 80:433-436(1999).
 DR EMBL; AF061755; AAD11961.1;
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein B.1.
 DR ProDom; PD000693; Glycoprot_B.1.
 SQ SEQUENCE 894 AA; 100415 MW; 806C1DD45B72AECB CRC64;

Query Match 40.7%; Score 48; DB 12; Length 894;
 Best Local Similarity 38.1%; Pred. No. 66;
 Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 2 IHHRGIVHAGRSIGRPLTG 22
 Db 721 IYSPFDGIGIGRAVGRVVTG 741

RESULT 14

O8UDH4 PRELIMINARY; PRT; 357 AA.
 AC O8UDH4;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Alcohol dehydrogenase.
 GN ADH OR ATU2151 OR AGR_C_3897.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. St.,
 RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
 RA Kuttyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Lao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RT C58.";
 RL Science 294:2317-2323(2001).

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughy J., Scott C., Lapps C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009165; AAL4141.1; ALT_INIT.
 DR EMBL; AE008131; AAK87898.1;
 KW Complete proteome.
 SQ SEQUENCE 357 AA; 38374 MW; 68F0CE54BD358020 CRC64;

Query Match 40.3%; Score 47.5; DB 16; Length 357;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

OY 4 HIRGIVHA-GRSIGRPLTG 22
 Db 70 HEFAGVISAVGRVTRKGTG 89

RESULT 15

O66090 PRELIMINARY; PRT; 273 AA.
 AC O66090;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Transmembrane protein Tmps (Fragment).
 OS Lactococcus lactis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1358;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MG1363;
 RX MEDLINE=98196737; PubMed=9537391;
 RA Poquet I., Ehrlich S.D., Gruss A.;
 RT "An export-specific reporter designed for gram-positive bacteria:
 application to Lactococcus lactis.";
 RL J. Bacteriol. 180:1904-1912(1998).
 DR EMBL; U95840; AAC14608.1;
 FT NON_TER 273
 SQ SEQUENCE 273 AA; 31411 MW; 64983CA4C554912E CRC64;

Query Match 39.8%; Score 47; DB 2; Length 273;
 Best Local Similarity 45.0%; Pred. No. 27;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 IHIFRGIVHAGSIGRELT 21
| | : : | | : | | | |
Db 51 IHSLYRNILHSGSGOGFLYT 70

Search completed: May 17, 2003, 14:27:12
Job time : 86 secs